



Db 466 GTNTYYTNTYNGGNGTNGCNGNWSNGCNCNGCNCARYTNTYNGCNTAYGARWSN 525  
QY 341 CAGCGACACGGGACATATATCCAGAAGGATTTCCTAGAGTCTATTACAACTCTGACCCCTG 400  
Db 526 MNGARTTYGAYGAYATHYTCARTGGGAYTTYACNGARGAYTTYTYAAYYTNACNYTN 585  
QY 401 AAGACCATGATGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGCGCGGTTTGTGATG 460  
Db 586 AARGARYTNCAYYTNCARMGNTGGTNGTNGCNGNTGYCCNCARGCNCAYTTYATGYTN 645  
QY 461 AAACAGACTCAGACATGTTTCATCAATTTGACTATCTGACTGAAGTCTTCTGAAAGAAA 520  
Db 646 AARGGNGAYGAYGAYGTTTGTNCAYGTNCCNAAAYGTYTNGARTTYTYTNGAYGGNTGG 705  
QY 521 AACAGAACACACAGGTTTTCACCTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGCAG 580  
Db 706 GAYCCNGCNCARGAYTYTNGTNGGNGAYGTNTHMGNCARGCNYTNCNAAAYMGNAAY 765  
QY 581 CCATTCAGCAAGTGGTTTGTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTC 640  
Db 766 ACNAAAGTNAARTAYTYVATHCCNCCNWSNATGTAYMNGNCNACNCAYTAYCCNCCNTAY 825  
QY 641 TGCTCCGGCACCGGCTACGTGTTTCTGGCGACGTGGCGAGTCAAGGTGTACAAATGTCCTC 700  
Db 826 GCNGGNGGNGGNGGNTAYGTNATGWSNMNGNCNACNGTNGMNGNYTNCARGCNAATHATG 885  
QY 701 AAGAGCGTCCCATACATTAACCTGGAAGACGTGTTTGTGGGGCTCTGCCCTCGAAAGGCTG 760  
Db 886 GARGAYGCNGARYTYTNWSNATHGAYGAYGTNTTYGTNGGNATGTGTYTNGMNGNYTN 945  
QY 761 AACATCAG 768  
Db 946 GGNYTWS 953

Search completed: April 11, 2003, 20:07:23  
Job time : 176.708 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 15:35:54 ; Search time 1036.47 Seconds  
(without alignments)  
14547.424 Million cell updates/sec

Title: US-09-914-152-2\_COPY\_401\_1331  
Perfect score: 931  
Sequence: 1 aatgggtttccgaagatga.....gggaagattgtccgcctgtc 931

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- EST::\*
- 1: em\_estba:\*\*
  - 2: em\_esthum:\*\*
  - 3: em\_estlin:\*\*
  - 4: em\_estmnu:\*\*
  - 5: em\_estov:\*\*
  - 6: em\_estpl:\*\*
  - 7: em\_estro:\*\*
  - 8: em\_htc:\*\*
  - 9: gb\_estli:\*\*
  - 10: gb\_est2:\*\*
  - 11: gb\_htc:\*\*
  - 12: gb\_est3:\*\*
  - 13: gb\_est4:\*\*
  - 14: gb\_est5:\*\*
  - 15: em\_estfun:\*\*
  - 16: em\_estom:\*\*
  - 17: gb\_gss:\*\*
  - 18: em\_gss\_hum:\*\*
  - 19: em\_gss\_inv:\*\*
  - 20: em\_gss\_pin:\*\*
  - 21: em\_gss\_vrt:\*\*
  - 22: em\_gss\_fun:\*\*
  - 23: em\_gss\_mam:\*\*
  - 24: em\_gss\_mus:\*\*
  - 25: em\_gss\_Other:\*\*
  - 26: em\_gss\_pro:\*\*
  - 27: em\_gss\_rod:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	366.6	39.4	434	12 BF748181	BF748181 RC3-BN041
C 2	326	35.0	338	9 AJ003597	AJ003597 AJ003597
C 3	234	25.1	400	17 A2768940	A2768940 LM0569K09
C 4	184	19.8	1101	17 CNS05B1	AL330238 Tetraodon
5	159.2	17.1	636	13 BG965086	BG965086 602829157
6	135	14.5	1514	11 AK003837	AK003837 Mus muscu

7	132.2	14.2	596	13	BJ501651	BJ501651
8	128.8	13.8	3564	11	BC028571	BC028571 Homo sapi
9	127	13.6	233	10	BB595953	BB595953
10	123.4	13.3	444	13	BM403751	BM403751 zam6155.2
11	122.6	13.2	879	14	BQ895238	BQ895238 AGENCOURT
12	117.2	12.6	965	17	CNS051B2	AL338663 Tetraodon
C 13	115.4	12.4	1021	17	CNS044T4	AL274513 Tetraodon
C 14	114.8	12.3	574	10	BE652985	BE652985 UT-M-AM1-
C 15	109.4	11.8	1101	13	BM548138	BM548138 AGENCOURT
16	108	11.6	620	9	AU133606	AU133606 AU133606
17	107.6	11.6	712	9	AL042887	AL042887 DKEZPA341
C 18	101.4	10.9	966	17	CNS04N8R	AL298404 Tetraodon
19	101	10.8	921	17	CNS03CTW	AL238253 Tetraodon
20	99.8	10.7	605	14	BQ417019	BQ417019 IK40901.Y
21	99.8	10.7	815	10	AW128625	AW128625 fe18C06.Y
C 22	97.2	10.4	464	9	AA972783	AA972783 oq01h06.S
23	95.6	10.3	611	13	BJ039496	BJ039496 BJ039496
C 24	95.4	10.2	411	9	AA858037	AA858037 ob13g10.S
25	93.8	10.1	793	12	BG207694	BG207694 RST27177
26	93.4	10.0	675	10	BB648067	BB648067 BB648067
C 27	92.6	9.9	809	17	CNS048G1	AL283131 Tetraodon
C 28	90.8	9.8	625	13	BJ054223	BJ054223 BJ054223
C 29	90.6	9.7	704	13	BM315483	BM315483 fw64h09.X
30	90.6	9.7	583	13	BI441305	BI441305 IC51h07.Y
31	87.6	9.4	1963	11	AK008674	AK008674 Mus muscu
32	83.6	9.0	923	9	AU067264	AU067264 AU067264
33	83.4	9.0	983	10	BB609433	BB609433 BB609433
34	82.8	8.9	584	10	BE373780	BE373780 601226187
35	80.2	8.6	360	9	AU171944	AU171944 AU171944
36	76.2	8.2	1041	17	CNS03BMU	AL236703 Tetraodon
37	70.6	7.6	427	13	BI476269	BI476269 fp54g03.Y
38	69.6	7.5	581	17	AZ381343	AZ381343 LM0137C22
C 39	68	7.3	1101	17	CNS05G62	AL335891 Tetraodon
40	68	7.3	1568	11	AK007600	AK007600 Mus muscu
41	66.4	7.1	675	12	BG077355	BG077355 H3014E06-
42	65.8	7.1	647	9	AL775073	AL775073 AL775073
C 43	65.6	7.0	909	9	AL522683	AL522683 AL522683
44	64.8	7.0	933	14	BQ926448	BQ926448 AGENCOURT
45	64.8	7.0	950	14	BQ946920	BQ946920 AGENCOURT

ALIGNMENTS

RESULT 1  
BF748181/c  
LOCUS BF748181 434 bp mRNA linear EST 10-JAN-2001  
DEFINITION RC3-BN0411-021000-021-c04 BN0411 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF748181  
VERSION BF748181.1 GI:12074857  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 434)  
AUTHORS Nagai M.A., da Silva W. Jr., Zago M.A., Bordin S., Costa F.F.,  
Dias Neto E., Garcia Correa R., Verjovski-Almeida S., Briones M.R.,  
Goldman G.H., Carvalho A.F., Matsumura A., Baia G.S., Simpson D.H.,  
Brunstein A., de Oliveira P.S., Bucher P., Jongeneel C.V., O'Hare  
M.J., Soares F., Brentani R.R., Reis L.F., de Souza S.J. and  
Simpson A.J.  
TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC3&t2=RC3-BN0411-021000-021-c04&t3=2000-10-02&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 24  
High quality sequence stop: 432.  
Location/Qualifiers  
1. 434

FEATURES  
source  
1. 434  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="BN0411"  
/dev\_stage="Adult"  
/note="Organ: breast\_normal; Vector: puc18; Site: 1; Sma1; Site: 2; Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
BASE COUNT 104 a 117 c 130 g 83 t  
ORIGIN

Query Match 39.4%; Score 366.6; DB 12; Length 434;  
Best Local Similarity 98.9%; Pred. No. 9.5e-97;  
Matches 369; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 559 CAATGAGTTTCCCATCAGGACGATTCAGCAAGTGTTTGTGACGTAATCTGAATATCC 618  
Db 434 CAATGAGTTTCCCATCAGGACGATTCAGCAAGTGTTTGTGACGTAATCTGAATATCC 375  
Qy 619 GTGGACAGGTACCCACCATTCGTCCGGACCGGCTAGCTGTTTTCGCGACGTGGC 678  
Db 374 GTGGACAGGTACCCACCATTCGTCCGGACCGGCTAGCTGTTTTCGCGACGTGGC 315  
Qy 679 GAGTCAGGTGTACATGTCTCCAGAGCGTCCCATACATTAACCTGGAAGACGTGTTGT 738  
Db 314 GAGTCAGGTGTACATGTCTCCAGAGCGTCCCATACATTAACCTGGAAGACGTGTTGT 255  
Qy 739 GGGGCTCTCCCTCGAAGCGTGAACATCAGATTGGAGAGCTCCACCTCCACGCGACCTT 798  
Db 254 GGGGCTCTCCCTCGAAGCGTGAACATCAGATTGGAGAGCTCCACCTCCACGCGACCTT 195  
Qy 799 TTTCAGGGGGCTTACGCTTCTCCGTATGCTTTCAGGAGGATCGTGGCTGCCACTT 858  
Db 194 TTTCAGGGGGCTTACGCTTCTCCGTATGCTTTCAGGAGGATCGTGGCTGCCACTT 135  
Qy 859 CATCAAGCTCGGACTCTCTTGACTACTGGCAGGCTCTAGAGAATTCCTCCGGGGGAAGA 918  
Db 134 CATCAAGCTCGGACTCTCTTGACTACTGGCAGGCTCTAGAGAATTCCTCCGGGGGAAGA 75  
Qy 919 TTGTCGGCTGTC 931  
Db 74 TTGTCGGCTGTC 62

RESULT 2  
AJ003597/c 338 bp mRNA linear EST 04-DEC-1997  
LOCUS  
DEFINITION  
AJ003597 Selected chromosome 21 cDNA library Homo sapiens cDNA  
clone MP19-12J9, mRNA sequence.  
ACCESSION  
AJ003597  
VERSION  
AJ003597.1 GI:2578270  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
1 (bases 1 to 338)  
Szulzewski, I., Hunt, E., Nguyen, M., Korn, B., Roehrdanz, B., Lehrach, H., and Vaspo, M.L.H.  
AUTHORS  
H. and Vaspo, M.L.H.  
TITLE  
An integrated transcript map for the whole human chromosome 21

Unpublished (1997)  
Contact: Vaspo, M.-L.  
Max Planck Institut fuer Molekulare Genetik  
Innestrasse 73, D14195 Berlin-Dahlem, Germany.  
Location/Qualifiers  
1. 338  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="21q"  
/clone\_lib="MP19-12J9"  
/note="Selected chromosome 21 cDNA library"  
/note="Swaroop et al. (1991) Nucleic Acids Res. 19: 1954."  
BASE COUNT 76 a 92 c 82 g 87 t 1 others  
ORIGIN

Query Match 35.0%; Score 326; DB 9; Length 338;  
Best Local Similarity 98.2%; Pred. No. 7.9e-85;  
Matches 329; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Qy 16 GATGAGATTGATGATATTGTCCTTCGTTCTGGGGCTCTTTGTTGTTATTTAGCAT 75  
Db 335 GACGAGATTGATGATATTGTCCTTCGTTCTGGGGCTCTTTGTTGTTATTTAGCAT 276  
Qy 76 GTACGCTCTAAATCCCTTCAAGAACAGTCTCTTTGTACAGAAAGAGCGGAACCTCT 135  
Db 275 GTACGCTCTAAATCCCTTCAAGAACAGTCTCTTTGTACAGAAAGAGCGGAACCTCT 216  
Qy 136 TAAGCTCCCATAGATACAGACTGCAGCAGACACCTCCCTTCCTCGTCTGCTGACCTC 195  
Db 215 TAAGCTCCCATAGATACAGACTGCAGCAGACACCTCCCTTCCTCGTCTGCTGACCTC 156  
Qy 196 ATCCCAACAAAGTGTGGCTGAGCGCATGGCCATCGGCGAGACGTGGGGAAGAGAGGAT 255  
Db 155 ATCCCAACAAAGTGTGGCTGAGCGCATGGCCATCGGCGAGACGTGGGGAAGAGAGGAC 96  
Qy 256 GGTGAAGGGAAGCAGCTGAAGCATTTCTCTCTGGGACACAGCAGTGCAGCGGA 315  
Db 95 GGTGAAGGGAAGCAGCTGAAGCATTTCTCTCTGGGACACAGCAGTGCAGCGGA 36  
Qy 316 AACGAAGAGGTGGACCAAGGAGAGCCAGCAGACAG 350  
Db 35 AACGAAGAGGTGGACCAAGGAGAGCCAGCAGCAGC 1

RESULT 3  
AZ768940/c 400 bp DNA linear GSS 16-FEB-2001  
LOCUS  
DEFINITION  
IM0569K09F Mouse 10kb plasmid UUC1M library Mus musculus genomic clone UUC1M0569K09 F, DNA sequence.  
ACCESSION  
AZ768940  
VERSION  
AZ768940.1 GI:12888559  
KEYWORDS  
GSS.  
SOURCE  
house mouse.  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 (bases 1 to 400)  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Rellly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A., and Wright, D., Weiss, R.  
AUTHORS  
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00





QY 582 CATTCAGCAGTGGTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATCT 641  
Db 246 TGGCGAGTAAAGTGGTACATTCGAGGACTTGTATCCGGAGAGCAATAACCCGCGTCT 187  
QY 642 GCTCCGGCAGCCGCTAGCTGTTTCTGGGACAGTGGGAGTCAGGTGTACATGTCTCCA 701  
Db 186 GCTCANGCACCNCCTAGCTGTTCTCAGCCGACGTGGCGGANGCTCATCTATAAGACCTCC 127  
QY 702 AGAGCTCCCATACATTAATTAACAGACAGCTGTTTGTGGGCTCTGCCCTCGAAGAGCTGA 761  
Db 126 TGCACACAGACTGTTGCACCTGGAGGACGTGTACNTGGGACTGTGTTTGCATAGCTGG 67  
QY 762 ACATC 766  
Db 66 GCATC 62  
  
RESULT 5  
BG965086  
LOCUS 602829157f1 NCI\_CGAP\_Co24 Mus musculus cDNA clone IMAGE:4983830 5',  
DEFINITION mRNA sequence.  
ACCESSION BG965086  
VERSION BG965086.1 GI:14352723  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 636)  
NIH-MGC http://mgc.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM10989 row: a column: 15  
High quality sequence start: 17  
High quality sequence stop: 634.  
FEATURES  
source Location/Qualifiers  
1..636  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4983830"  
/clone\_lib="NCI\_CGAP\_Co24"  
/lab\_host="DH10B (TI phage-resistant)"  
/notes="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.6 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."  
BASE COUNT 155 a 172 c 155 g 154 t  
ORIGIN  
Query Match 17.1%; Score 159.2; DB 13; Length 636;  
Best Local Similarity 66.3%; Pred. No. 1.4e-35;  
Matches 277; Conservative 0; Mismatches 133; Indels 8; Gaps 3;  
  
QY 2 ATGGCTTCCGAGATGAGATGATGATATTTGCTTCCTGTTCTGGGCGCTTTTCT 61  
Db 225 ATGGCTCACATGAGACAGGCTAGTTTACGCTCCCTCCATCTCATGATGGCGCACTCTGC 284  
QY 62 TGTCTATTTAGCATGTACATCTAAATCCTTTCAGAGACACTCTTGTTCACAGAAA 121  
Db 285 TTGCTACTTCAGCATGGATCTTTTCA-----GAGAACTCCCGTGTGTTTAAACGAAAGT 338

QY 122 GACGGGAACCTTCTTAAGCTCCAGATAC-AGACTGCAGGCAGACACCTCCCTTCTCTCGT 180  
Db 339 CACGGGAAGTCTCTTACAGATCCGATATAGACTGCAAGCAGACGCGCTTCTCTGGT 398  
QY 181 CTTGCTGGTACCTCATCCCAACAGTTGGCTGAGCGATGCGCATCGGCGAGAGCTG 240  
Db 399 GCTGCTGGTACGTCATCTCACAAGCAGCTGGCGGCTCGCATGGCCATCCGCAAGAGCTG 458  
QY 241 GGGGAAGAGAGAGGATGGTGAAGGG-AAAGCAGCTGAAGACATCTTCTTCTCTGGGACCA 299  
Db 459 GGGTAGAGAGACATCTGTGAGGGGACACAGGTGAGGACCTTCTTCTTCTGGGACCT 518  
QY 300 CCAGCAGTGGCAGCGAAAGAGAGGTGGACAGGAGCGGACCGACGACATTA 359  
Db 519 CCGACAGCAGGAGAGATGGACGCCAACACCTTGAGAGCGAGCAGCAGCGACATTA 578  
QY 360 TCCAGAAGGATTTCTAGACGCTCTATTACAATCTGACCTCGAAGACCATGATGGGCAT 417  
Db 579 TCCAGAGGACTTCAAGGATGCTTAACCTGACCTGACCTGAGACCATGATGGGTAT 636  
  
RESULT 6  
AK003837  
LOCUS 1514 bp mRNA linear HTC 19-JAN-2002  
DEFINITION Mus musculus 18 days embryo whole body cDNA, RIKEN full-length  
enriched library, clone:1110020A06:UDP-Gal:betaGal:AC beta  
1,3-galactosyltransferase, polypeptide 3, full insert sequence.  
ACCESSION AK003837  
VERSION AK003837.1 GI:12834746  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (strain:C57BL/6J) 18 days embryo cDNA to mRNA,  
clone\_lib:RIKEN full-length enriched mouse cDNA library  
clone:1110020A06.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1  
REFERENCE  
AUTHORS Carninci,P. and Hayashizaki,Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PUBMED 10349636  
REFERENCE 2  
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,  
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE 20499374  
PUBMED 11042159  
REFERENCE 3  
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,  
Konno,H., Akiyama,J., Nishi,K., Kusunagi,T., Tashiro,H., Itoh,M.,  
Lumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,  
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,  
Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,T., Watanabe,M.,  
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,  
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
TITLE RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multipillar sequencer  
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
MEDLINE 20530913  
PUBMED 11076861  
REFERENCE 4  
AUTHORS Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,  
Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,  
Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,  
Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,  
Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,  
Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,  
Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,  
Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,  
Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,



KEYWORDS EST. Japanese medaka.  
SOURCE ORGANISM Oryzias latipes  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;  
Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.  
REFERENCE 1 (bases 1 to 596)  
AUTHORS Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.  
TITLE Medaka EST Project in Takeda's lab  
JOURNAL Unpublished (2001)  
COMMENT Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.  
FEATURES  
source  
1. 596  
/organism="Oryzias latipes"  
/strain="d-rp"  
/db\_xref="taxon:8090"  
/clone\_lib="MF01FSA CDNA"  
/sex="mixture of female and male"  
/tissue\_type="whole embryo"  
/dev\_stage="fry stage 40"  
BASE COUNT 138 a 168 c 151 g 137 t 2 others  
ORIGIN  
Query Match 14.2%; Score 132.2; DB 13; Length 596;  
Best Local Similarity 55.2%; Pred. No. 1.2e-27;  
Matches 277; Conservative 0; Mismatches 222; Indels 3; Gaps 1;  
QY 268 GCAGCTGAAGACATCTTCCTCTGGGACACAGAGCTGAGCGGGAAGCAAGGT 327  
Db 13 GCGAGTGTGTCACCCCTTTCTTGTGGGCGCAGCATGGACGTGTCTCTGAACCATGTT 72  
QY 328 GGACGAGGAGCGGACAGCGGGACATATCCAGAGGATTTCTAGACGTCTATTA 387  
Db 73 GGACGAGGAGTGCAGATCTTTCATGTAGTGTGGGAGACTTTATTGACTCGTACCA 132  
QY 388 CAATCTGACCTGAAGACCATGATGGGATAGATGGGTCCATCGCTTTGTCTCAGGC 447  
Db 133 CAACCTGACGCTNAAGACCCCTGATGGGTATGGCTGGTGGCTACTGACCAAGGC 192  
QY 448 GCGCTTTGTGATGAACAGACTCAGACATGTCATCAATGTTGACTATCTGACTGAAC 507  
Db 193 CCAATATGCTCAAGACGGACAGTGACATCTTTGTCAACATGGAGAACCTCATCAAA 252  
QY 508 GCTTCTGAG--AAAACAGAACACAGGTTTTCACCTGGCTTCTTGAACCTCAATGA 564  
Db 253 CTTTCTGAAGCCCAACCAAGCCTTAGGAGAAGGATTTTACAGGTTACGTCAATCAACGG 312  
QY 565 GTTTCCTCAGGAGCCATTCAGCAAGTGGTTTGTGAGTAAATCTGAATCTCGTGGGA 624  
Db 313 GGGCCGATCAGACATACAGCAAGTGGTACATGCCCGGATCTTTATCCGACAG 372  
QY 625 CAGGTACCCACCATCTCTCGCGGACCGGCTACGTTGTTTTCGGGACGAGTGGCGAGTCA 684  
Db 373 CAATATCCCACTTCTGCTCGGACCGGCTACGTTCTCTCAGCAGATGTAGCTGAGCT 432  
QY 685 GGTGTACAAATGCTCCAGAGCGTCCCATACATTAACTTGAAGACAGTGTGTTGGGGCT 744  
Db 433 CATATACNAGATCTCTTTGACACACAGGCTCTCCACCTGGGAGGAGCTGTATGAGT 492  
QY 745 CTGCTCGAAGGCTGAACATC 766  
Db 493 TTGCTCGTGAAGCTGGGCATC 514

RESULT 8

BC028571 3564 bp mRNA linear HTC 21-AUG-2002  
LOCUS Homo sapiens, UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase,  
polypeptide 3, clone IMAGE:4838965, mRNA.  
DEFINITION BC028571  
ACCESSION BC028571.1 GI:22382221  
VERSION HTC.  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 3564)  
AUTHORS Strausberg,R.  
TITLE Direct Submission  
JOURNAL Submitted (23-APR-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov  
COMMENT Contact: MGC help desk  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Institute for Systems Biology  
http://www.systemsbio.org  
contact: amadan@systemsbiology.org  
Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha  
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Series: IRAC Plate: 34 Row: p Column: 23  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 15451873  
This clone has the following problem: incomplete processing.  
FEATURES  
Location/Qualifiers  
1. 3564  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4838965"  
/tissue\_type="Testis"  
/clone\_lib="NIH\_MGC\_97"  
/lab\_host="DH10B"  
/note="Vector: pBluescript"  
BASE COUNT 1048 a 661 c 741 g 1114 t  
ORIGIN  
Query Match 13.8%; Score 128.8; DB 11; Length 3564;  
Best Local Similarity 48.7%; Pred. No. 2.4e-26;  
Matches 386; Conservative 0; Mismatches 397; Indels 9; Gaps 1;  
QY 111 TTACAAGAAAGACGGGAACCTTCCTTAAGCTCCAGATACAGACTGAGGAGACACCTC 170  
Db 864 TTACAGACAGACTTTCACCTTCACACTTCGAGAGCATTCACAACTGCTCTCAAAATC 923  
QY 171 CCTTCTCGTCTGCTGGTGACCTCATCCACAAACAGTTGGCTGAGCGCATGCCATCC 230  
Db 924 TATTCTTGCTCATCTCTGGTGACCTCCACCTTCAGATGTGAAGCCAGCGGACCATTA 983  
QY 231 GCGACAGTGGGGAAAGAGAGGATGGTGAAGGAAAGACAGCTGAAGACATCTTTCCTCC 290  
Db 984 GAGTTACTTGGGGTGAAGAAAGAGTCTTTGGGGGATATGAGTTCTTACATTTTCTTAT 1043  
QY 291 TGGGGACCCACAGCAGCTGCGCGGAAACGAAAGAGG-----TGGACGAGGAGGCC 341  
Db 1044 TAGGCCAAGAGGCTGAAGAGGAGACAAATGTTGGCATTGCTCTAGAGGATGAACACC 1103  
QY 342 AGCGACACGGGACATATTCCAGAGGATTTCTCTAGACGCTATTACAATCTGACCTGA 401  
Db 1104 TTCTTTATGGTGACATAATCCGACAAAGATTTTGTAGACATATAATAACCTGACCTTGA 1163



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.  
1 (bases 1 to 444)  
Ton, C., Dempsey, A.A., Hwang, D.M. and Liew, C.C.  
Identification and Characterization of Expressed Sequence Tags from Zebrafish Skeletal Muscles cDNA Library  
Unpublished (2002)  
Contact: Liew CC  
Brigham and Women's Hospital  
Harvard Medical School  
75 Francis St. Boston, MA 02115, USA  
Tel: 6177328915  
Fax: 6179750995  
Email: cliew@rics.bwh.harvard.edu  
PCR Primers  
FORWARD: 5' GCCAGCTCGAAATTAACCTTCACCTAAAGGG 3'  
BACKWARD: 5' CCAGTGAATTGTATAGGACTCACTATAGGGCG 3'  
Seq primer: T3.

Location/Qualifiers  
1. 444  
/organism="Danio rerio"  
/db\_xref="taxon:7955"  
/clone="zam6155"  
/dev\_stage="adult"  
/clone\_lib="zebrafish skeletal muscles cDNA library"  
/note="Organ: skeletal muscles; Vector: Lambda ZAP Express; Site\_1: EcoRI; Site\_2: XhoI; mRNA was purified from zebrafish skeletal muscles. cDNA was synthesized using a XhoI-oligo dt adaptor-primer. EcoRI adaptors were ligated, followed by digestion with XhoI, for directional cloning into pre-digested lambda ZAP Express vector."

BASE COUNT 120 a 115 c 102 g 105 t 2 others

Query Match 13.3%; Score 123.4; DB 13; Length 444;  
Best Local Similarity 60.7%; Pred. No. 4.3e-25;  
Matches 238; Conservative 0; Mismatches 147; Indels 7; Gaps 2;

QY 326 GTGACGAGGAGCGACGACGAGGACATATCCAGAGGATTCCTAGACGCTAT 385  
DB 39 GTTGATGAGGAGCTTACAGCACCACATCATCCAGCAGGACTTCTTGACACCTAT 98  
QY 386 TACAATCTGACCTGAGACCATGATGGGCTAGAAATGGTCCATCGCTTTTGTCTCAG 445  
DB 99 TACAACCTTACTATTAAGACCTAATGGCATGAGTGGTTCCTGCTGCTATTCGCCCTC 158  
QY 446 GCGCGTTTGTGATGAACACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTG 505  
DB 159 GCTGGTTATGTCATGAAGCAGGACGACATGTTTGCAACACCGGAATATCTTATCCAA 218  
QY 506 CTGCTTCTGAAGAAAAACAGAACAAAC-----CAGGTTTTTCTGCTGGCTTCTTGA 559  
DB 219 AAGTGTCTAAGCCAAACACGCGACCTCGACAGAACTTACTTCACAGGCTACCTAATG 278  
QY 560 AATGAGTTTCCATCAGCAGCCATTCAGCAAGTGGTTGTGTCAGTAAATCTGAATATCC 619  
DB 279 GGCTACGCTCCCAACCGCAACAGGACAGTAAGTGGTATCATGCCACGAGCTGTATCC 338  
QY 620 TGGACAGGTACCCACCATTTCTGTCGCGCAGCGCTAGCTGTTTTCTGGGACGTTGG 679  
DB 339 AGTGAGATACCCCTATCTTTGTCGCGGACAGGTTATGTTCTCGGGGACATGGCA 398  
QY 680 AG-TCAGGTGTACAAATGCTCCAGAGCGTCC 710  
DB 399 NGCGAAATCTACAATGCCTCATTGAGTATCC 430

RESULT 11  
BO895238  
LOCUS BO895238 879 bp mRNA linear EST 16-AUG-2002  
DEFINITION AGENCOURT\_8747587 NIH\_MGC\_130 Mus musculus cDNA clone IMAGE:6332353 5', mRNA sequence.

ACCESSION BO895238  
VERSION BO895238.1 GI:22287252  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 879)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: rga@p5-r@mail.nih.gov  
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.

cDNA Library Preparation: ResGen, Invitrogen Corp  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM13789 row: n column: 02  
High quality sequence stop: 547.

FEATURES  
source Location/Qualifiers  
1..879  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6332353"  
/clone\_lib="NIH\_MGC\_130"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: otocysts; Vector: PCMV-SPORT6.1.ccd; Site\_1: EcoRV; Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.95 kb. Constructed by ResGen, Invitrogen Corp. Note: this is a NIH\_MGC Library."

BASE COUNT 236 a 213 c 206 g 224 t

ORIGIN

Query Match 13.2%; Score 122.6; DB 14; Length 879;  
Best Local Similarity 52.0%; Pred. No. 9.7e-25;  
Matches 275; Conservative 0; Mismatches 254; Indels 0; Gaps 0;

QY 327 TGGACGAGGAGCGACGACGAGGACATTCACAGAGGATTCCTAGACGCTATT 386  
DB 112 TGGAGGATGAGCAGCGTCTCTATGGTGATATTATACGGCAAGACTTTCTAGACACAT 171  
QY 387 ACAATCTCACCTGAAGACCATGATGGGATAGAAATGGTCCATCGCTTTTGTCTCAGG 446  
DB 172 ATAATTGACCTTGAACACCATTTATGGCTTCAGGTGGGTAATGGAGTTTGGCCCAATG 231  
QY 447 CGGGTTTTGTGATGAACACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGA 506  
DB 232 CCAAGTATATCATGAACACAGACATGATGTTTCATCAACACTGGCAATTTAGTCAAGT 291  
QY 507 TGCCTTCTGAAGAAACAGACAAACAGGTTTTTCTACTGGCTTCTTGAACACTCAATGAGT 566  
DB 292 ATCTTTTAAACCTAAACCACTCAGAGAAGTTTTTTCACGGGCTATCTCTAATGATAACT 351  
QY 567 TTCCCATCAGCAGCCATTCAGCAAGTGGTTGTCTAGTAAATCTGAATATCCGTGGGACA 626  
DB 352 ATTCTCTAGAGGATTTTTCATATAAACACCATTTTCATACCAAGAGTACCCCTTCAAGG 411  
QY 627 GGTACCCACCATTCGCTCCGGCACCGCTACGTGTTTCTGGGACGTGGCGAGTCAAG 686  
DB 412 TGTTCCTCCCTACTGCAGCGGGCTGGGTACATTATGTCGGGAGCTGGTGGCCAGGG 471  
QY 687 TGTACAATGTCTCCAGAGCGTCCCATACATTAACCTGGAAGACGTTGTGTGGGGTCT 746  
DB 472 TCTACGAGATGATGAGTCACGTGAAGCCCATCATCAAGTTTGAAGAGTTTATGTTGGCATCT 531  
QY 747 GCCTCGAAGGCTGAACATCATGATGGAGAGCTCCACCTCCCGACCGACCTTTTTCAG 806  
DB 532 GCTTGAATTTGCTCAAGTGGGACATTCATATCCAGAAGACACCAACCTTTTCTTCTGT 591

```
QY 807 GGGGTTACCGCTTCGGTATGCGTTCCTTCAGGAGGATCGTGCGCTGCCA 855
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 592 ACAGATCCACTTGGATGATGTCAGCTCAGACGCGTGATTCGACGCCA 640

RESULT 12
CNS05IB2      CNS05IB2      965 bp      DNA      linear      GSS 26-MAY-2000
LOCUS      Tetraodon nigroviridis genome survey sequence T3 end of clone
DEFINITION      045P20 of library A from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION      AL338663
VERSION      AL338663.1 GI:8232421
KEYWORDS      GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
ORGANISM      Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE      1 (bases 1 to 965)
AUTHORS      Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
      Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
      Saurin,W. and Weissenbach,J.
      Human gene number estimate provided by genome wide analysis using
      Tetraodon nigroviridis DNA sequence
      Unpublished
JOURNAL      2 (bases 1 to 965)
REFERENCE      Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
AUTHORS      Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
      Weissenbach,J.
      Characterization and repeat analysis of the compact genome of the
      freshwater pufferfish Tetraodon nigroviridis
      Unpublished
JOURNAL      3 (bases 1 to 965)
REFERENCE      Genoscope.
AUTHORS      Direct Submission
TITLE      Submitted (12-APR-2000)
JOURNAL      This sequence is a single read and was generated as part of a large
COMMENT      scale clone-end sequencing project of the Tetraodon nigroviridis
      genome. For more information, please take a look at
      http://www.genoscope.cns.fr/Tetraodon.
      Location/Qualifiers
FEATURES      source
      1. 965
      /organism="Tetraodon nigroviridis"
      /db_xref="taxon:99883"
      /clone_lib="A"
      /clone="045P20"
      /note="Genoscope sequence ID : C0AA045DH10A1-end : T3"
BASE COUNT      249 a 245 c 220 g 229 t 22 others
ORIGIN

Query Match      12.6%; Score 117.2; DB 17; Length 965;
Best Local Similarity      54.6%; Pred. No. 3.9e-23;
Matches 242; Conservative      0; Mismatches 199; Indels      2; Gaps      1;

QY 325 GTGGACGAGGAGACGACGACGACGGGACATATCCAGAGGATTTCCTAGAGCTCTA 384
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 42 GCTGGACGAGGAGATCGATCTTCCACACATCGTTGTGGAGGATTTCGACTCGTA 101
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 385 TTACAATCTGACCTGAGACCATGATGGCGATGAGATGGTCCATCGCTTTTCTCTCA 444
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102 CCACAATTGACACTCAAGACCCCTGATGGCATGGCGTGGCTACGTTTCTCTCAA 161
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 445 GCGCGGTTTGTGTGAAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGA 504
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 162 GGCTCAGTATGTCNCTAAAGACGACAGTACATCTTTGTCAACATGGAATCTCATCTT 221
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 505 ACTGCTTCTGAGAAAAACAGACACACAG - GTTTTCACTGGCTTCTTGAACTCAAT 562
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 222 TAACCTCTGAAACCCCAACACCAACCCAGAGGAGGTACTTCACTGTTATGATNATCAA 281
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 563 GAGTTTCCCATCAGGCAGCCATTTCAGCAAGTGGTTTTCAGTAATCTGAATATCCGTGG 622
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 282 TGGTGGGCCAATNAGAGACATGCGCANAAGTGGTACATGTCNAGAGATCTGTACCCAGAG 341
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 623 GACAGGTACCCACCATTTCTCGCGCACCGGTACGTGTTTCTCGGCGAGTGCGCGAGT 682
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 342 AGCAAGTNCGCCCTTTGTTCCGGCACTGGTACGTCTTCTCCGCGGAGTGCGCCGAG 401
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 683 CAGGTGTACAATGTCTCCAGAGAGCGTCCCATACATAATTAACATGGAAGAGCTGTTGTGGG 742
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 402 CTATATTTTATNCTCTTACACACAGCGTCTGCACCTGGAGGACGTGTATNTGGGT 461
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 743 CTCTGCTCGAAAGGCTGAACAT 765
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 462 NTGTGCTTCGCAAGCTGGGCAT 484
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
CNS044T4/c      1021 bp      DNA      linear      GSS 18-MAY-2000
LOCUS      Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
DEFINITION      082009 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION      AL274513
VERSION      AL274513.1 GI:7996991
KEYWORDS      GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
ORGANISM      Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE      1 (bases 1 to 1021)
AUTHORS      Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
      Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
      Saurin,W. and Weissenbach,J.
      Human gene number estimate provided by genome wide analysis using
      Tetraodon nigroviridis DNA sequence
      Unpublished
JOURNAL      2 (bases 1 to 1021)
REFERENCE      Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
AUTHORS      Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
      Weissenbach,J.
      Characterization and repeat analysis of the compact genome of the
      freshwater pufferfish Tetraodon nigroviridis
      Unpublished
JOURNAL      3 (bases 1 to 1021)
REFERENCE      Genoscope.
AUTHORS      Direct Submission
TITLE      Submitted (12-APR-2000)
JOURNAL      This sequence is a single read and was generated as part of a large
COMMENT      scale clone-end sequencing project of the Tetraodon nigroviridis
      genome. For more information, please take a look at
      http://www.genoscope.cns.fr/Tetraodon.
      Location/Qualifiers
FEATURES      source
      1. 1021
      /organism="Tetraodon nigroviridis"
      /db_xref="taxon:99883"
      /clone="082009"
      /clone_lib="G"
      /note="Genoscope sequence ID : C0BG082AH05SP1-end :
      PUC-Ori"
BASE COUNT      207 a 262 c 290 g 259 t 3 others
ORIGIN

Query Match      12.4%; Score 115.4; DB 17; Length 1021;
Best Local Similarity      59.6%; Pred. No. 1.4e-22;
Matches 214; Conservative      1; Mismatches 138; Indels      6; Gaps      1;

QY 153 ACTGCGAGGAGACACCTCCCTTCCTCGTGGTGGTGAACCTCATCCCAACAGTTGG 212
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 365 AGTGCAGGAGAGACTCCCTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 306
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

QY	213	CTGAGCGCATGTCCATCCGGCAGACGTGGGGGAAGAAGAGAGTAGTGTGAAGGAAACGCAGC	272
Db	305	ATCCGCGTAACGCCGTCATCCGCGCACACCGTTGGGGAACAAGACAGACAGCGATGGGCGCTGGGCT	246
QY	273	TGAAGACATTCTTCCTCCTCGTGCGGACCACCAAGCAGTCAGTCAGCGGAAACGAAAAGAGG-----	326
Db	245	TCGTTTGGCCTTCTTCCTGCTCGGAACGGGGAGGAAGTCGGACACTTTTCTCCAANGAGCA	186
QY	327	TGGACCAAGSAGAGCCAGCGACACGGGACATAATTCCAGAAGGATTTCCCTAGACGTCTATT	386
Db	185	TCGAGGAAGAGAGGCCAATTTACCATGCATCATCCAACAGGACTACCAAAGACACCTACT	126
QY	387	ACAACTGACCCCTGAAGACACCATGATGGGCATAGATGGGTCCATCGCTTTTGTCCTCAGG	446
Db	125	ACAACTGACCATTAACAACCTGATGGGTGAAGTCTGAGCTGGGTGGCCACCTATTGGCCCACG	66
QY	447	CGGGCTTTGTGATGAAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAA	505
Db	65	CCTCTACGTGATGAAAAACAGACGACGACATGTTGTCAACACAGAGTATCHCAACCA	7
RESULT	14		
BE652985/c			
LOCUS	BE652985	574 bp mRNA linear EST 06-SEP-2000	
DEFINITION	UI-M-AMI-afz-e-01-0-UI.r1 NIH_BMAP_MAM_N Mus musculus cDNA clone UI-M-AMI-afz-e-01-0-UI 5', mRNA sequence.		
ACCESSION	BBE652985		
VERSION	BBE652985.1 GI:9978848		
KEYWORDS	EST,		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 574)		
AUTHORS	Bonaldo,M.F., Lennon,G. and Soares,M.B.		
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery		
JOURNAL	Genome Res. 6 (9), 791-806 (1996)		
MEDLINE	97044477		
COMMENT	Contact: Chin, H National Institute of Mental Health 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD 20892-9643, USA Tel: 301 443 1706 Fax: 301 443 9890 Email: mBEST@mail.nih.gov cDNA Library Preparation: M.B. Soares Lab Clone distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements Seq primer: M13 Reverse.		
FEATURES	Location/Qualifiers		
source	1..574 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="UI-M-AMI-afz-e-01-0-UI" /clone_lib="NIH-BMAP_MAM_N" /dev_stage="27-32 days" /lab_host="DRI0B (Life Technologies)" /note="vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; The NIH_BMAP_MAM_N library is a normalized library constructed from mouse amygdala. The tag is a string of 5 nucleotides present between the Not I site and the oligo-dT track. The library was constructed as described by Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996. Tissue provided by Ms. Annie Novakovich, Zivic-Miller Laboratories."		
BASE COUNT	161 a	113 c	121 g
ORIGIN	178 t	1 others	



High quality sequence stop: 418.  
Location/Qualifiers  
1..1101  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5733782"  
/clone\_lib="NIH\_MGC\_124"  
/tissue\_type="hippocampus"  
/lab\_host="DH10B"  
/note="Organ: Brain; Vector: pCMV-SPORT6; Site.1: EcoRV (destroyed); Site.2: NotI; RNA source male hippocampus, age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 0.9-4 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 012."  
BASE COUNT 304 a 239 c 238 g 319 t 1 others  
ORIGIN  
Query Match 11.8%; Score 109.4; DB 13; Length 1101;  
Best Local Similarity 51.4%; Pred. No. 8.1e-21;  
Matches 285; Conservative 0; Mismatches 261; Indels 9; Gaps 1;  
QY 111 TTACAGAAACACGGGAACCTTCCTTAAGCTCCAGATACAGACTGCAGGAGACACCTC 170  
DB 222 TTACAGACAAGACTTTCACCTTCAGACTTCGAGAGCATTCAAACTGCTCTCATCAAAATC 281  
QY 171 CCTTCTCTGCTGCTGGTGACCTCATCCACAAACAGTTGGCTGAGCGCATGGCCATCC 230  
DB 282 CATTTCTGCTATCTTGGTGACCTCCACCCTTCAGATGTGAAGCCAGCGAGGCCATTA 341  
QY 231 GCGAGAGTGGGGAAGAGAGGATGTTGAAGGAAAGAGAGCTGAAGAGACATTTCTTCCTCC 290  
DB 342 GAGTTACTTGGGTGAAAAAGTCTTGGTGGGATATGAGGTTCTTACATTTTCTTAT 401  
QY 291 TGGGGACCAACAGCAGTGCAGCGGAACAGAAAGAGG-----TGGACCAGGAGAGCC 341  
DB 402 TAGGCCAAGAGGCTGAAAGAGAGACAAATGTTGGCATTTGCTTACAGGATGAACACC 461  
QY 342 AGCGACACGGGACATATCCAGAAGATTTCCTAGAGCTCTATTACAACTGACCCCTGA 401  
DB 462 TTCTTTATGGTGACATAATCCGACAGATTTTATAGACACATATAATAACCTGACCTTGA 521  
QY 402 AGACCATGATGGCATAGATGGTCCATCGCTTTTGTCTCCTCAGCGGCGTTTGTGATGA 461  
DB 522 CAACCATATTGGCATTACAGTGGGTAACTGAGTTTGGCCCAATGCCAAGTACGTAATGA 581  
QY 462 AAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAATGCTTCTGAAGAAA 521  
DB 582 AGACACACACTGATGTTTTCATCAATGCTGCAATTTAGTGAAGTATCTTTTAAACCTAA 641  
QY 522 ACAGAAACACAGGTTTTTTCAGTGGTCTCTTGAACCTCAATGAGTTTCCCATCAGGACG 581  
DB 642 ACCACTCAGAGAAGTTTTTTCACAGGTTATCTCTAAATTGATAATTATTCTTATAGAGGAT 701  
QY 582 CATTACAGAGTGGTTGTGTCAGTAAATCTGAATATCCGTGGGAGAGGTACCCACCATCT 641  
DB 702 TTACCAAAAAACCCATATTCTTACAGGAGTATCCTTTCAAGGCGTTCCTCCATCACT 761  
QY 642 GCTCCGGCACCGGCT 656  
DB 762 GCAGCGGTAGGGGT 776

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 15:37:18 ; Search time 25.7081 Seconds  
(without alignments)  
11106.080 Million cell updates/sec

Title: US-09-914-152-2\_COPY\_401\_1331  
Perfect score: 931  
Sequence: 1 aatgggttccgaagatga.....gggaagattgtccgcgtgc 931

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114.8	12.3	1134	4	US-09-482-180A-3
2	78.4	8.4	1420	4	US-09-482-180A-1
3	68.6	7.4	1191	4	US-09-459-133-3
4	67	7.2	1167	4	US-09-459-133-14
5	54.4	5.8	1434	2	US-09-055-097-2
6	41.6	4.5	1532	4	US-09-459-133-1
7	41.4	4.4	7218	1	US-08-232-463-14
8	40.2	4.3	1446	4	US-09-459-133-12
9	37.4	4.0	6503	4	US-09-404-650-12
10	36.6	3.9	289	4	US-09-007-005-17
11	36.6	3.9	289	4	US-09-244-796-17
12	35	3.8	2557	4	US-08-464-954A-1
13	34.6	3.7	248	4	US-09-007-005-32
14	34.6	3.7	248	4	US-09-244-796-32
15	34.6	3.7	277	4	US-09-007-005-3
16	34.6	3.7	277	4	US-09-244-796-3
17	34.2	3.7	1100	3	US-09-248-335-53
18	34.2	3.7	1241	1	US-08-471-033-39
19	34.2	3.7	1241	1	US-08-471-033-42
20	34.2	3.7	1241	2	US-08-471-044-39
21	34.2	3.7	1241	2	US-08-471-044-42
22	34.2	3.7	1241	2	US-08-463-483A-39
23	34.2	3.7	1241	2	US-08-463-483A-42
24	34.2	3.7	1241	2	US-08-471-046A-39
25	34.2	3.7	1241	2	US-08-471-046A-42
26	34.2	3.7	1241	2	US-08-470-566B-39
27	34.2	3.7	1241	2	US-08-470-566B-42

28	34.2	3.7	1241	2	US-08-469-334-39	Sequence 39, Appl
29	34.2	3.7	1241	2	US-08-469-334-42	Sequence 42, Appl
30	34.2	3.7	1241	3	US-09-300-529-39	Sequence 39, Appl
31	34.2	3.7	1241	3	US-09-300-529-42	Sequence 42, Appl
32	34.2	3.7	1358	1	US-08-471-033-45	Sequence 45, Appl
33	34.2	3.7	1358	2	US-08-471-044-45	Sequence 45, Appl
34	34.2	3.7	1358	2	US-08-463-483A-45	Sequence 45, Appl
35	34.2	3.7	1358	2	US-08-471-046A-45	Sequence 45, Appl
36	34.2	3.7	1358	2	US-08-470-566B-45	Sequence 45, Appl
37	34.2	3.7	1358	3	US-08-469-334-45	Sequence 45, Appl
38	34.2	3.7	1358	3	US-09-300-529-45	Sequence 45, Appl
39	34.2	3.7	1389	1	US-08-471-033-27	Sequence 27, Appl
40	34.2	3.7	1389	2	US-08-471-044-27	Sequence 27, Appl
41	34.2	3.7	1389	2	US-08-463-483A-27	Sequence 27, Appl
42	34.2	3.7	1389	2	US-08-471-046A-27	Sequence 27, Appl
43	34.2	3.7	1389	2	US-08-470-566B-27	Sequence 27, Appl
44	34.2	3.7	1389	2	US-08-469-334-27	Sequence 27, Appl
45	34.2	3.7	1389	3	US-09-300-529-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1  
US-09-482-180A-3  
; Sequence 3, Application US/09482180A  
; Patent No. 6361985  
; GENERAL INFORMATION:  
; APPLICANT: Conklin, Darrell C.  
; APPLICANT: Yamamoto, Gayle  
; APPLICANT: Gao, Zeren  
; APPLICANT: Whitmore, Theodore E.  
; APPLICANT: Jaspers, Stephen  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING  
; TITLE OF INVENTION: BETA-1,3-GALACTOSYLTRANSFERASE HOMOLOG, 2NSSP6  
; FILE REFERENCE: 98-80  
; CURRENT APPLICATION NUMBER: US/09/482,180A  
; CURRENT FILING DATE: 2000-01-12  
; PRIOR APPLICATION NUMBER: 60/115,721  
; PRIOR FILING DATE: 1999-01-12  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 1134  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: degenerate sequence  
; NAME/KEY: variation  
; LOCATION: (1)...(1134)  
; OTHER INFORMATION: n is any nucleotide  
; NAME/KEY: misc.feature  
; LOCATION: (1)...(1134)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-482-180A-3

Query Match	12.3%	Score 114.8;	DB 4;	Length 1134;
Best Local Similarity	31.3%;	Pred. No. 6.1e-26;		
Matches 209;	Conservative 98;	Mismatches 358;	Indels 3;	Gaps 1;
QY	104	TCCTTTGTTTACAGAAAGACGGGAACCTCTTAAAGTCCACAGATACAGACTGCAGGCAG	163	
DB	286	TTTNTACNTATGNCAYTGYMGNAAYTTTWSNATHYTNNGARCCNWSNGGTGYWSN	345	
QY	164	ACACCTCCCTTCTCTGCTGCTGGTGACCTCATCCACAAACAGTTGGCTGAGGCATG	223	
DB	346	AARGAYACNTTYYTNTYNGCNATHAARWSNCARCCNCGNCAYGTNGARMNGNNGCN	405	
QY	224	GCATCCCGCAGACGTGGGGA---AAGAGAGGATGTGAAGGAAAGCAGCTGAAGACA	280	
DB	406	GCNATHMGWSNACNTGGGNGMNGTNGGNGTGGCNGMNGMNCARYTNARVTN	465	
QY	281	TTCTTCTCTCTGGGGACCAACAGCAGTGCAGCGGAAAGAGGTGGACCGAGGAGC	340	



```

QY 170 CCCTTCCTCGTCTGCTGACCTCATCCACAAACAGTTGGCTGAGCGCATGCCATC 229
  || || || || || || || || || || || || || || || || || || || ||
Db 442 CCNTAYTNTYNTGNCNGTNAARWSNGARCCNGNMGNTTYGCGARWNGCARGCNGTN 501
  || || || || || || || || || || || || || || || || || || || ||
QY 230 CGGCAGAGTGGGGAAAGAGAGATGGTGAAGGAAAGAGCTGAAGACATCTTCCTC 289
  || || || || || || || || || || || || || || || || || || || ||
Db 502 MNGARACNTGGGNGWNSCNCNGCNGGNTAHMGNNTYNTTYTNTYNTGNSNCCN 561
  || || || || || || || || || || || || || || || || || || || ||
QY 290 CTGGGACACACAGCAGTGCAGCGGAAAGAAAGAGTGGACAGAGAGCGAGCGACAC 349
  || || || || || || || || || || || || || || || || || || || ||
Db 562 GTNGGARGCNGCNGCNGAYYNTGAYWS---NYTNGTNGCNGTGGGARWSNMGNTAY 618
  || || || || || || || || || || || || || || || || || || || ||
QY 350 GGGGACATATCAGAGAGATTCCTAGACGCTCTATTACAATCTCACCTGAGACCATG 409
  || || || || || || || || || || || || || || || || || || || ||
Db 619 WSNAGAYTNTYNTGCGAYTNTYNTGAYTNCNTTYAAYCARACNTYNAARGAYTN 678
  || || || || || || || || || || || || || || || || || || || ||
QY 410 ATGGGATCAATAGGTCCTCGCTTTTGTCTCAGCGGGCTTTGTGATGAAACAGAC 469
  || || || || || || || || || || || || || || || || || || || ||
Db 679 YTYNTYNTGNTGGYNTGGNMGNCAYTYCCNACNGNWSNTTYTNTYNTGNGCNCAR 738
  || || || || || || || || || || || || || || || || || || || ||
QY 470 TCAGACATGTTCAATATGTTGACTATCTGACTGAATGCT-----TCTGAAGAAAC 523
  || || || || || || || || || || || || || || || || || || || ||
Db 739 GAYGAYCCTTYTNCAYACNCCNGCNYTNTGNCNAYYTNMGNCNYTNCNCCNGCN 798
  || || || || || || || || || || || || || || || || || || || ||
QY 524 AGAACACACAGGTTTTCACGTGCTTCTTGAAACTCAATGAGTTTCCCATCAGCGAGCCA 583
  || || || || || || || || || || || || || || || || || || || ||
Db 799 WSNCGNMGNSNYTNTAYTNGGARGTNTTYACNARGCNCATGCCNTNMGNAARCCN 858
  || || || || || || || || || || || || || || || || || || || ||
QY 584 TTAGCAAGTGGTTTTCAGTAATCTGAATATCCGTGGGACAGGTACCCACATCTGC 643
  || || || || || || || || || || || || || || || || || || || ||
Db 859 GNGGNCNTTYTAYTNCNARGNSNTTYTYGARGNG---GNTAYCNCNGNTAYGNC 915
  || || || || || || || || || || || || || || || || || || || ||
QY 644 TCGGACCGGCTACGTGTTTTCGGGACGTCGGGAGTGCAGTGTACATGTCTCCAAG 703
  || || || || || || || || || || || || || || || || || || || ||
Db 916 WSNGGNGGNGTAYGNTNATHGCGNGNMGNTYNTGNCNCCNTGGYTYNTNMGNGCNGCN 975
  || || || || || || || || || || || || || || || || || || || ||
QY 704 AGCTCCCATACATTAACCTGGAAGACGTTTGTGGGCTGCTGCTGAAAGCGTGAAC 763
  || || || || || || || || || || || || || || || || || || || ||
Db 976 MNGTNCNCCNTTYCCNTTYGARGAYTNTAYACNGGNTYNTGYATHMGNGCNYTNGN 1035
  || || || || || || || || || || || || || || || || || || || ||
QY 764 ATCAGATTGAGAGGCTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCTCC 823
  || || || || || || || || || || || || || || || || || || || ||
Db 1036 YTGNTCCNARGCNCAYCCNGGNTTYTNCNGCNGTGGCNGCNGAYMGNACNGCNGAY 1095
  || || || || || || || || || || || || || || || || || || || ||
QY 824 GTATGCTCTTCAGAGGATCGTGGCCTGCCACTTCATCAAGCCTCGGACTCTCTGGAC 883
  || || || || || || || || || || || || || || || || || || || ||
Db 1096 CAYTGCTNTYMGNAAYTYNTYNTGNTNMGNCNYTNGNCCNARGCNCNWSNATHMGN 1155
  || || || || || || || || || || || || || || || || || || || ||
QY 884 TACTGGGAGGCTCTAGAGAAATTCGG 909
  || || || || || || || || || || || || || || || || || || || ||
Db 1156 YNTGGAARCAARYTNCARGAYCCNMG 1181
  || || || || || || || || || || || || || || || || || || || ||

```

RESULT 4  
 US-09-459-133-14  
 ; Sequence 14, Application US/09459133  
 ; Patent No. 6415988  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Conklin, Darrell C.  
 ; APPLICANT: Yamamoto, Gayle  
 ; APPLICANT: Jaspers, Stephen R.  
 ; APPLICANT: Gao, Zeren  
 ; TITLE OF INVENTION: BETA-1,3-GALACTOSYLTRANSFERASE HOMOLOGS  
 ; FILE REFERENCE: 98-77  
 ; CURRENT APPLICATION NUMBER: US/09/459,133  
 ; CURRENT FILING DATE: 1999-12-10  
 ; PRIOR APPLICATION NUMBER: 60/111,697  
 ; PRIOR FILING DATE: 1998-12-10  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 14  
 ; LENGTH: 1167  
 ; TYPE: DNA

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: degenerate sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1167)
; OTHER INFORMATION: n = A,T,C or G
US-09-459-133-14

Query Match          7.2%; Score 67; DB 4; Length 1167;
Best Local Similarity 31.1%; Pred. No. 5.4e-11;
Matches 183; Conservative 74; Mismatches 319; Indels 12; Gaps 3;

QY 170 CCCTTCCTCGTCTGCTGACCTCATCCACAAACAGTTGGCTGAGCGCATGCCATC 229
  || || || || || || || || || || || || || || || || || || || ||
Db 418 CCNTAYTNTYNTGNCNGTNAARWSNGARCCNGGNCAYTYTTCGCGARWNGCARGCNGTN 477
  || || || || || || || || || || || || || || || || || || || ||
QY 230 CGGCAGAGTGGGGAAAGAGAGATGGTGAAGGAAAGAGCTGAAGACATCTTCCTC 289
  || || || || || || || || || || || || || || || || || || || ||
Db 478 MNGARACNTGGGNGWNSCNCNGTNGCNGNMGNYTNTTYTNTYNTGNGWNSCNCN 537
  || || || || || || || || || || || || || || || || || || || ||
QY 290 CTGGGACACACAGCAGTGCAGCGGAAAGAAAGAGTGGACAGAGAGCGAGCGACAC 349
  || || || || || || || || || || || || || || || || || || || ||
Db 538 YTNNG---NATGGGNGCNCNAYYTNMGNSNYTNGTNCNTGGGARWSNMGNTAY 594
  || || || || || || || || || || || || || || || || || || || ||
QY 350 GGGGACATATCCAGAAGGATTCCTAGACGCTCTATTACAATCTGACCCCTGAGACCATG 409
  || || || || || || || || || || || || || || || || || || || ||
Db 595 GNGAYTNTYNTTGGGAYTNTYNTGAYTNCNTAYAYMGNACNYTNAARGAYTN 654
  || || || || || || || || || || || || || || || || || || || ||
QY 410 ATGGGATCAATAGGTCCTCGCTTTTGTCTCAGCGGGCTTTGTGATGAAACAGAC 469
  || || || || || || || || || || || || || || || || || || || ||
Db 655 YTYNTYNTNACGTGTYTNSNCAYCATGYCCNGAYGTNAAYTYTGTNTYNTNARGTNCAR 714
  || || || || || || || || || || || || || || || || || || || ||
QY 470 TCAGACATGTTCAATATGTTGACTATCTGACTGAATGCTTCTGAA-----GAAAAAC 523
  || || || || || || || || || || || || || || || || || || || ||
Db 715 GAYGAYCCTTYTNCAYATHCCNCGCNYTNGARCAAYTNCARACNYTNCNCCNACN 774
  || || || || || || || || || || || || || || || || || || || ||
QY 524 AGAACACACAGGTTTTCACGTGCTTCTTGAAACTCAATGAGTTTCCCATCAGCGAGCCA 583
  || || || || || || || || || || || || || || || || || || || ||
Db 775 TGGGNCNMGNSNYTNTAYTNGGNGARATHYTYACNARGCNCNARCCNNTNMGNAARCCN 834
  || || || || || || || || || || || || || || || || || || || ||
QY 584 TTCACAAAGTGGTTTGTCAATATCAATATCCGTGGGACAGGTACCCACCATTTCTGC 643
  || || || || || || || || || || || || || || || || || || || ||
Db 835 GNGGNCNCCNTTYTAYTNCNNAARACNTTYTYGARGNGAY---TAYCCNCCNTAYGNC 891
  || || || || || || || || || || || || || || || || || || || ||
QY 644 TCGGACACGCGCTACGTTTTCGCGAGTGGCGAGTGCAGGTGTACAATGTCTCCAAG 703
  || || || || || || || || || || || || || || || || || || || ||
Db 892 WSNGGNGGNGTAYGNTNATHWSNMGNYTNGCNCNTGGYTYNTNARGCNGCNGCN 951
  || || || || || || || || || || || || || || || || || || || ||
QY 704 AGCTCCCATACATTAACCTGGAAGACGTTTGTGGGCTCTGCTC 751
  || || || || || || || || || || || || || || || || || || || ||
Db 952 MNGTNGCNCNTTYCCNTTYGAYGAYGNTAYACNGGNTTYTGTYT 999
  || || || || || || || || || || || || || || || || || || || ||

```

RESULT 5  
 US-09-055-097-2  
 ; Sequence 2, Application US/09055097  
 ; Patent No. 5955282  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hillman, Jennifer L.  
 ; APPLICANT: Guegler, Karl J.  
 ; APPLICANT: Corley, Neil C.  
 ; APPLICANT: Shah, Purvi  
 ; APPLICANT: Patterson, Chandra  
 ; TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 ; STREET: 3174 Porter Drive  
 ; CITY: Palo Alto  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94304

```

; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows version 3.0
; SEQ ID NO 1
; LENGTH: 1532
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: {101}...(1294)
US-09-459-133-1

Query Match      4.5%; Score 41.6; DB 4; Length 1532;
Best Local Similarity 48.6%; Pred. No. 0.0054;
Matches 144; Conservative 0; Mismatches 149; Indels 3; Gaps

QY 170 CCCTTCTGCTGCTGGTGACCTCATCCACAAACAGTTGGCTAGCGCATGCCATC 229
      |||||
Db 542 CCCTAGCTGCTGTTGGCCGCTCAAGTCAGAACACAGGCGGCTTTGCAGAACGACAGCGCGTG 601
      |||||

QY 230 CGCAGACAGCTGGGGGAAAGAGAGGATGGTGAAGGGAAAGCAGCTGAAGACATCTTCCCTC 289
      |||||
Db 602 AGAGACAGTGGGCGAGTCCAGCTCCAGGATCCGGCTGCTCTTCTGCTAGGGTCTCCG 661

QY 290 CTGGGACACACAGCAGTGCAGCGGAAACGAAGAGAGTGACACGAGAGCAGCGACAC 349
      |||||
Db 662 GTGGGTGAGCGGGGGGCTGACCTAGACTC---ACTAGTGCATGGGAGAGCGCTGCTGCTAC 718

QY 350 GGGGACATTATCCAGAAGATTTTCCTAGAGCTCTATTACAATCTGACCCCTGAAGACCATG 409
      |||||
Db 719 AGTGACCTGCTGCTGGGACTTCCTCGAGGTCCTCCATTCAACAGACGCTCAAGACCTG 778

QY 410 ATGGGCATAGAATGGGTCCATCGCTTTTGTTCCTACGGCGGGTTTGTGATGAAC 465
      |||||
Db 779 CTGCTGCTGGCCCTGGCTGGCGGCCCACTGCCCAACCGTGAAGTTTGTCTTGGGAGC 834
      |||||

```

[illegible]

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZgt-F1s  
US-08-232-463-14

Query Match 4.4%; Score 41.4; DB 1; Length 7218;  
Best Local Similarity 9.0%; Pred. No. 0.015;  
Matches 36; Conservative 187; Mismatches 178; Indels 0; Gaps 0;  
QY 135 TTAAGTCCAGATACAGACTGAGGACACACCTCCTCTCTGCTGCTGCTGCTGCT 194  
DB 1516 TTAAGTTTCAAAAACGGCATGTAGGCATCACTGTAAATTACCTATCTATGAAGTAGTTA 1457  
QY 195 CATCCCAACAAACAGTTGGCTGAGCGCATGCGCATCCGCGACAGCTGGGGGAAAGAGAGGA 254  
DB 1456 AAGATAGAAGAAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1397  
QY 255 TGGTGAAGGAAAGCAGCTGAAGCATTTCTCTCTGGGACACACAGCAGTGCAGGG 314  
DB 1396 RRR 1337  
QY 315 AAACGAAAGGTGGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 374  
DB 1336 RRR 1277  
QY 375 TAGACCTCTATTACATCTGACCTGAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 434  
DB 1276 RRR 1217  
QY 435 TTGTCTCAGGGCGGTTTGTGATGAAACAGACTCAGACATGTCATCAATGTGACT 494  
DB 1216 RRR 1157  
QY 495 ATCTGCTCAACTGCTCTGTAAGAAACAGAAACAGAAACAGAAACAGAAACAGG 535  
DB 1156 RRR 1116

RESULT 8  
US-09-459-133-12  
; Sequence 12, Application US/09459133  
; Patent No. 6416988  
; GENERAL INFORMATION:  
; APPLICANT: Conklin, Darrell C.  
; APPLICANT: Yamamoto, Gayle  
; APPLICANT: Jaspers, Stephen R.  
; APPLICANT: Gao, Zeren  
; TITLE OF INVENTION: BETA-1,3-GALACTOSYLTRANSFERASE HOMOLOGS  
; FILE REFERENCE: 98-77  
; CURRENT APPLICATION NUMBER: US/09/459,133  
; PRIOR FILING DATE: 1999-12-10  
; PRIOR FILING DATE: 60/411,697  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 12  
; LENGTH: 1446  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (107)...(1273)  
US-09-459-133-12

Query Match 4.3%; Score 40.2; DB 4; Length 1446;

Best Local Similarity 48.5%; Pred. No. 0.014;  
Matches 142; Conservative 0; Mismatches 148; Indels 3; Gaps 1;  
QY 170 CCCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 229  
DB 524 CCCTACTTGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 583  
QY 230 CGGACAGCTGGGGGAAAGAGAGATGTTGAAGGAAAGACAGCTGAAGACATTTCTTCTCTC 289  
DB 584 AGGAGACCTGGGGCAGCCAGTTGCTGGGACCCCGTTGCTCTCTCTGCTGGGTCCTCC 643  
QY 290 CTGGGGACACACAGCTGCGAGCGGAAAGAGAGGTGACAGGAGGAGGAGGAGGAGGAGGAGGAG 349  
DB 644 CTAGT---AATGGGGGGCTGACTTAAGATCAGCTGGTGACGTGGGAAAGCCCGCGTAT 700  
QY 350 GGGACATTTATCCAGAAGATTTCTAGACGTCTATTACAATCTGACCTTGAAGACCATG 409  
DB 701 GGTGACCTACTGCTGCTGGGACTTCTCTGATGTTCTCTACACCGGACACTCAAGGACCTG 760  
QY 410 ATGGGCATAGATAGGTCCATCGCTTTTGTCTCTAGGCGGCGTTTGTGATGAA 462  
DB 761 CTGCTGCTGACCTGCTGAGCCACCACTGCCCGCATGTCAATTTTCTCTGCA 813

RESULT 9  
US-09-404-650-12/c  
; Sequence 12, Application US/09404650  
; Patent No. 6309858  
; GENERAL INFORMATION:  
; APPLICANT: Dietrich, Paul S.  
; APPLICANT: McGovern, Joseph G.  
; TITLE OF INVENTION: T-TYPE CALCIUM CHANNEL VARIANTS; COMPOSITIONS THEREOF;  
; TITLE OF INVENTION: AND USES  
; FILE REFERENCE: R0043B-REG sequence listing  
; CURRENT APPLICATION NUMBER: US/09/404,650  
; CURRENT FILING DATE: 1999-09-23  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 6503  
; TYPE: DNA  
; ORGANISM: Rattus sp.  
US-09-404-650-12

Query Match 4.0%; Score 37.4; DB 4; Length 6503;  
Best Local Similarity 53.8%; Pred. No. 0.25;  
Matches 77; Conservative 0; Mismatches 66; Indels 0; Gaps 0;  
QY 152 GACTGCGAGGACAGACACCTCCCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 211  
DB 4197 GCGCCACGCGAGTCAGATCGGTGGTGTGTTCTCCACTCCCAACAAATGGTAG 4138  
QY 212 GCTGAGCGCATGGCCATCCGGCAGACGTGGGGGAAAGAGAGGATGTTGAAGGAAAGCAG 271  
DB 4137 AACTTGCTTTGAAAGCTGCACCCCGAGGATGCCGAAGATGATGAAGAGGCACAGCAG 4078  
QY 272 CTGAAGACATTTCTCTCTCTGGG 294  
DB 4077 ATGAGGACGATGTTCCCAATGGG 4055

RESULT 10  
US-09-007-005-17  
; Sequence 17, Application US/09007005B  
; Patent No. 6258558  
; GENERAL INFORMATION:  
; APPLICANT: Szostak, Jack W.  
; APPLICANT: Roberts, Richard W.  
; APPLICANT: Liu, Rihe  
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN  
; TITLE OF INVENTION: FUSIONS  
; FILE REFERENCE: 00786/350003  
; CURRENT APPLICATION NUMBER: US/09/007,005B

```
; CURRENT FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; NAME/KEY: misc_feature
; LOCATION: (1)..(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

Query Match          3.9%; Score 36.6; DB 4; Length 289;
Best Local Similarity 9.7%; Pred.No. 0.078;
Matches 28; Conservative 109; Mismatches 152; Indels 0; Gaps 0;

QY    240 CGGGGAAGACAGATCGTCAAGGAAAGCAGCTGAAGACATTTCCTCCTGGGACCA 299
      :|::||:|::||:|::||:|::||:|::||:|::||:|::||:|::||:|::||:|:
DB     1 RGRGRARCARURUARCRRURRURARURURARCRARURURARCRARURGRNRN 60
      ::::::::::::::::::::

QY    300 CCAGCAGTGCCAGCGGAAACAAGAGGTGCACCAGGAGGCCAGCGACGGCACATTA 359
      :|::||:|::||:|::||:|::||:|::||:|::||:|::||:|::||:|::||:|:
DB     61 RSNNRNRSNNRSNNRSNNRSNNRSNNRSNNRSNNRSNNRSNNRSNNRSNNRN 120
      ::::::::::::::::::::

QY    360 TCCAGAAGSGATTTCTTAGACGCTATTACAACTCATGCCCTGAAGACCATGATGGCATA 419
      :|::||:|::||:|::||:|::||:|::||:|::||:|::||:|::||:|::||:~
DB     121 RSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRN 180
      ::::::::::::::::::::

QY    420 AATGGGTCCAATCGCTTTTGCTTCAGCGCGGCTTGTTGATGAAAACAGACTCACACATGT 479
      :|::||:|::||:|::||:|::||:|::||:|::||:|::||:|::||:|::||:~
DB     181 RSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRNRNSRRCRGRCRGRRARA 240:
      ::::::::::::::::::::

QY    480 TCATCAATGTTGACTATCTGACTGAAC TGCTTCCTGAGAAAAAACAAGAAC 528
      :|::||:|::||:|::||:|::||:|::||:|::||:|::||:|::||:|::||:~
DB     241 RCRRRCRRURGRGRCRUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCC 289
      ::::::::::::::::::::

RESULT 11
US-09-244-796-17
; Sequence 17, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Bihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: 00786/350007
; CURRENT APPLICATION NUMBER: US/09/244,796
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; EARLIER APPLICATION NUMBER: 09/007,005
; EARLIER FILING DATE: 1998-01-14
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; NAME/KEY: misc_feature
```

```
; ANTI-SENSE: NO
US-09-464-954A-1

Query Match      3.8%; Score 35; DB 4; Length 2557;
Best Local Similarity 48.7%; Pred. No. 0.83;
Matches 95; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

Qy 508 GCTTCTGAGAGAAACAGAACACACAGGTTTTTCACTGGCTTTTGAACATCAATGAGTT 567
      || ||||| || || || || || || || || || || || || || || || || ||
Db 865 GCGCTGGAGAGAGAGAGATCAAAAGAGAGAAAGGAGTCCATGCGCCCTCAATGAGAA 924

Qy 568 TCCATCAGGAGCCATTCAGCAAGTGGTTTTCAGTAATCTGAATATCGCTGGGACAG 627
      || || || || || || || || || || || || || || || || || || || ||
Db 925 GCAGATCCTCGAGAGTCAACAGTTCAGTTTGGTCAACCTGGCCCTATGCTACGAGAC 984

Qy 628 GTACCCACCATCTGCTCGCGCACCGGCTACGTTTCTGGCAGCTGGCGAGTCAAGT 687
      || || || || || || || || || || || || || || || || || || || ||
Db 985 CAACGATGCATCTGCTTGGTCTGACCATCATGATGGGGGTGACGTGAAGTTCCACAT 1044

Qy 688 GTACAAATGCTCCAA 702
      ||||| |||||
Db 1045 CTACAACATGGGCAA 1059

RESULT 13
US-09-007-005-32
; Sequence 32, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; EARLIER FILING DATE: 1998-01-14
; EARLIER FILING DATE: 1997-01-27
; EARLIER FILING DATE: 1997-01-27
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-007-005-32

Query Match      3.7%; Score 34.6; DB 4; Length 248;
Best Local Similarity 19.2%; Pred. No. 0.3;
Matches 41; Conservative 78; Mismatches 94; Indels 0; Gaps 0;

; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; EARLIER FILING DATE: 1998-01-14
; EARLIER FILING DATE: 1997-01-27
; EARLIER FILING DATE: 1997-01-27
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-007-005-32

Query Match      3.7%; Score 34.6; DB 4; Length 248;
Best Local Similarity 19.2%; Pred. No. 0.3;
Matches 41; Conservative 78; Mismatches 94; Indels 0; Gaps 0;

Qy 236 ACGTGGGGAAGAGAGAGATGGTGAAGGAAAGAGAGCTGAAGACATTTCTCTCTGGGG 295
      ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 35 RCRARURURARCRARURGRGRCRURGRARARARARARARARARARARARARCRURGRA 94

Qy 296 ACCACAGCAGTGCAGCGGAAAGAGAGAGTGGACGAGGAGGAGGAGGAGGAGGAGG 355
      ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 95 RUCRUCRURGRARARARARARARARARARARARARARARARARARARARARARCRURC 154

Qy 356 ATTATCCAGAGGATTTCTTCTAGCTCTATTACATCTGACCTGAAGACCATGATGGGC 415
      ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 155 RGRURGRARARARARARARARARARARARARARARARARARARARARARARARCRURC 214

Qy 416 ATAGATGGGTCCATCGCTTTTGTCTCAGCGG 448
      ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 215 RURGRGRURARARARARARARARARARARARARARARARARARARARARARARCRURC 247

RESULT 15
US-09-007-005-3
; Sequence 3, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; EARLIER FILING DATE: 1998-01-14
; EARLIER FILING DATE: 1997-01-27
; EARLIER FILING DATE: 1997-01-27
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
US-09-007-005-3

Query Match      3.7%; Score 34.6; DB 4; Length 277;
Best Local Similarity 19.2%; Pred. No. 0.3;
Matches 41; Conservative 78; Mismatches 94; Indels 0; Gaps 0;

Qy 236 ACGTGGGGAAGAGAGAGATGGTGAAGGAAAGAGAGCTGAAGACATTTCTCTCTGGGG 295
      ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 35 RCRARURURARCRARURGRGRCRURGRARARARARARARARARARARARARARCRURGRA 94

Qy 296 ACCACAGCAGTGCAGCGGAAAGAGAGTGGACGAGGAGGAGGAGGAGGAGGAGGAGG 355
      ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 95 RUCRUCRURGRARARARARARARARARARARARARARARARARARARARARARCRURC 154

Qy 356 ATTATCCAGAGGATTTCTTCTAGCTCTATTACATCTGACCTGAAGACCATGATGGGC 415
      ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 155 RGRURGRARARARARARARARARARARARARARARARARARARARARARARARCRURC 214

Qy 416 ATAGATGGGTCCATCGCTTTTGTCTCAGCGG 448
      ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 215 RURGRGRURARARARARARARARARARARARARARARARARARARARARARARCRURC 247

RESULT 14
US-09-244-796-32
; Sequence 32, Application US/09244796
```



```
Best Local Similarity 19.28; Pred. No. 0.32;
Matches 41; Conservative 78; Mismatches 94; Indels 0; Gaps 0;

QY 236 ACGTGGGGAAAGAGAGATGGTGAAGGGAAGCAGCTGAAGACATTTCTCTCTGGGG 295
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 35 RCRARURURARCRARURGRGRCRURGRARARGRARGRARGRARGRARGRARGR 94
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 296 ACCACCAGCAGTCAGCGGGAACGAAGAGGTGGACCAGGAGCCAGCACGGGGAC 355
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 95 RURCRURGRARARGRARGRARGRARGRARGRARGRARGRARGRARGRARGRURC 154
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 356 ATTATCCAGAGGATTTCTTAGAGCTCTATTACAATCTGACCTGAAGACCATGATGGC 415
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 155 RGRURGRARARGRARGRARGRARGRARGRARGRARGRARGRARGRARGRARGRC 214
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 416 ATAGATGGTCCATCGGTTTTTGCTCTCAGGCG 448
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 215 RURGRCRURARARGRARGRARGRARGRARGRARGRARGRARGRARGRARGRCR 247
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: April 11, 2003, 22:41:42  
Job time : 41.7081 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 14:22:18 ; Search time 1693.98 Seconds  
(without alignments)  
15994.713 Million cell updates/sec

Title: US-09-914-152-2\_COPY\_401\_1331

Perfect score: 931

Sequence: 1 aatggctttccgaagatga.....gggaagattgtccgcctgtc 931

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*

1: gb.ba:\*

2: gb.htg:\*

3: gb.in:\*

4: gb.om:\*

5: gb.ov:\*

6: gb.pat:\*

7: gb.ph:\*

8: gb.pl:\*

9: gb.pr:\*

10: gb.ro:\*

11: gb.sts:\*

12: gb.sy:\*

13: gb.un:\*

14: gb.vi:\*

15: em.ba:\*

16: em.fun:\*

17: em.hum:\*

18: em.in:\*

19: em.mu:\*

20: em.om:\*

21: em.or:\*

22: em.ov:\*

23: em.pat:\*

24: em.ph:\*

25: em.pl:\*

26: em.ro:\*

27: em.sts:\*

28: em.un:\*

29: em.vi:\*

30: em.htg.hum:\*

31: em.htg.inv:\*

32: em.htg.other:\*

33: em.htg.mus:\*

34: em.htg.pln:\*

35: em.htg.rod:\*

36: em.htg.man:\*

37: em.htg.virt:\*

38: em.sy:\*

39: em.htgo.hum:\*

40: em.htgo.mus:\*

41: em.htgo.other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	931	100.0	2762	9	AB020337	AB020337 Homo sapi
2	931	100.0	2775	6	E38419	E38419 Novel poly
3	930	99.9	933	9	AF145784	AF145784 Homo sapi
4	930	99.9	10562	6	E38420	E38420 Novel poly
5	930	99.9	170121	9	AF064860	AF064860 Homo sapi
6	930	99.9	340000	9	HS21C080	AL163280 Homo sapi
7	921	98.9	933	9	HS46078	AJ006078 Homo sapi
8	898.8	96.5	1576	9	AB041416	AB041416 Homo sapi
9	882.2	94.8	1570	9	AB041415	AB041415 Pan panis
10	879.2	94.4	1565	9	AB041413	AB041413 Homo sapi
11	872	93.7	1360	9	AB041412	AB041412 Gorilla g
12	869.2	93.4	1566	9	AB041414	AB041414 Pan trogl
13	850	91.3	1579	9	AB041417	AB041417 Pongo pyg
14	517.8	55.6	927	10	AF254738	AF254738 Mus muscu
15	517.8	55.6	196900	2	AC020851	AC020851 Mus muscu
16	513	55.1	175861	2	AC120346	AC120346 Mus muscu
17	320.8	34.5	65400	2	AC109264	AC109264 Mus muscu
18	185.2	19.9	917	10	AB039137	AB039137 Mus muscu
19	185.2	19.9	917	10	AB039140	AB039140 Mus muscu
20	185.2	19.9	917	10	AB039141	AB039141 Mus muscu
21	185.2	19.9	917	10	AB039142	AB039142 Mus muscu
22	183.6	19.7	917	10	AB039136	AB039136 Mus muscu
23	183.6	19.7	917	10	AB039138	AB039138 Mus muscu
24	183.6	19.7	917	10	AB039143	AB039143 Mus spici
25	182	19.5	917	10	AB039134	AB039134 Mus muscu
26	182	19.5	917	10	AB039135	AB039135 Mus muscu
27	182	19.5	917	10	AB039139	AB039139 Mus muscu
28	182	19.5	1175	10	AF029790	AF029790 Mus muscu
29	182	19.5	219368	10	AC098721	AC098721 Mus muscu
30	166	17.8	1053	9	AB041409	AB041409 Pan panis
31	165	17.7	1056	9	AB041410	AB041410 Gorilla g
32	165	17.7	1059	9	AB041411	AB041411 Pongo pyg
33	164.4	17.7	1060	9	AB041408	AB041408 Pan trogl
34	163.4	17.6	1037	9	AB041407	AB041407 Homo sapi
35	163.4	17.6	1739	6	E07739	E07739 cDNA encodi
36	163.4	17.6	202001	9	AC016723	AC016723 Homo sapi
37	161.8	17.4	1909	9	AF117222	AF117222 Homo sapi
38	142.8	15.3	996	10	AB039156	AB039156 Mus muscu
39	142.8	15.3	996	10	AB039158	AB039158 Mus muscu
40	142.8	15.3	996	10	AB039162	AB039162 Mus muscu
41	142.8	15.3	996	10	AB039163	AB039163 Mus spici
42	141.2	15.2	996	10	AB039154	AB039154 Mus muscu
43	141.2	15.2	996	10	AB039155	AB039155 Mus muscu
44	141.2	15.2	996	10	AB039157	AB039157 Mus muscu
45	141.2	15.2	996	10	AB039159	AB039159 Mus muscu

#### ALIGNMENTS

RESULT 1  
AB020337  
LOCUS  
DEFINITION Homo sapiens mRNA for UDP-Gal:GlcNAc betal,3-galactosyltransferase 5, complete cds.  
ACCESSION AB020337  
VERSION AB020337.1 GI:4835502  
KEYWORDS UDP-Gal:GlcNAc betal,3-galactosyltransferase 5.  
SOURCE Homo sapiens Adenocarcinoma cell\_line:Colo 205 cDNA to mRNA.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (sites)  
AUTHORS Isshiki,S., Togayachi,A., Kudo,T., Nishihara,S., Watanabe,M., Kubota,T., Kitajima,M., Shiraishi,N., Sasaki,K., Andoh,T. and

TITLE		Cloning, expression, and characterization of a novel UDP-galactose:beta-N-acetylglucosamine beta1,3-galactosyltransferase (beta3Gal-T5) responsible for synthesis of type 1 chain in colorectal and pancreatic epithelia and tumor cells derived therefrom	
JOURNAL	J. Biol. Chem. 274 (18), 12499-12507 (1999)		
MEDLINE	99230269		
REFERENCE	2 (bases 1 to 2762)		
AUTHORS	Isshiki, S., Togayachi, A. and Narimatsu, H.		
TITLE	Direct Submission		
JOURNAL	Submitted (20-NOV-1998) Hisashi Narimatsu, Soka University, Institute of Life Science; 1-236, Tangei-cho, Hachioji, Tokyo 192-8577, Japan [E-mail: sissihiki@po.iiijnet.or.jp, Tel:81-426-91-9466, Fax:81-426-91-9315]		
FEATURES	Location/Qualifiers		
source	1..2762		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="21"		
	/map="21q22.3"		
	/cell_line="Colo 205"		
	/cell_type="Adenocarcinoma"		
gene	1..2762		
	/gene="beta1,3-GalT 5"		
exon	1..273		
	/gene="beta1,3-GalT 5"		
	/number=1		
exon	274..433		
	/gene="beta1,3-GalT 5"		
	/number=3		
exon	434..2762		
	/gene="beta1,3-GalT 5"		
	/number=4		
CDS	434..1366		
	/gene="beta1,3-GalT 5"		
	/codon_start=1		
	/product="UDP-Gal:GlcNAc beta1,3-galactosyltransferase 5"		
	/protein_id="BAA77664.1"		
	/db_xref="GI:4835503"		
	/translation="MAPPKRLMYICLLVLGALCLYFSMYSLNPFKEQSFVYKKDNFS LKLPDQCTPFVLLVSHKQAERMAIQWKGKMGVGRKOLKTFLLDLGITSS AAETKEVDSEGRHGDIOKDFLDVYNTLKTMMGIEWHFRFCPOAFAVMTDSDF INVDYITELLKKNTTRFTGFLKLNFFIRQPSKVSFVSEYEPWDRYPFPCSGTG VYFGSDVAGQVNVSKSPYIKLEDVFVGLCLERINIRLEELHSQPTFFPGGLRFSVC LFRRIVACHFIKPTLLDYWQALENSRGDCPPV"		
polyA_site	2762		
	/gene="beta1,3-GalT 5"		
	/note="45 a nucleotides"		
BASE COUNT	651 a 703 c 679 g 729 t		
ORIGIN			
Query Match	100.0%; Score 931; DB 9; Length 2762;		
Best Local Similarity	100.0%; Pred. No. 1.6e-242;		
Matches 931; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 AATGGCTTCCCGAAGATGAGATTGATGTATATTGGCTTCTGGTCTGGGGGCTCTTTG 60		
Db	433 AATGGCTTCCCGAAGATGAGATTGATGTATATTGGCTTCTGGTCTGGGGGCTCTTTG 492		
QY	61 TTGTGATTTTAGCATGTACAGTCTAAATCCCTTTCAAAGAACAGTCCTTTGTTTACAAGAA 120		
Db	493 TTTGTATTTTAGCATGTACAGTCTAAATCCCTTTCAAAGAACAGTCCTTTGTTTACAAGAA 552		
QY	121 AGACGGGAATCTCCTTAAGCTCCAGATACAGACTGAGCGACAGACACCTCCCTTCCTCGT 180		
Db	553 AGACGGGAATCTCCTTAAGCTCCAGATACAGACTGAGCGACAGACACCTCCCTTCCTCGT 612		
QY	181 CCTGCTGTCACCTCATCCCAACAAACAGTTGGGTGAGCGCATGCCATCGCCACAGACGCTG 240		
Db	613 CCTGCTGTCACCTCATCCCAACAAACAGTTGGGTGAGCGCATGCCATCGCCACAGACGCTG 672		
QY	241 GGGGAAAGAGAGGATGGTGAAGGGAAAGACAGCTGAAGACATTTCTTCCTCGTGGGGACAC 300		

PC C12P21/08,C12Q1/68,G01N33/53/(C12N1/21,C12R1:185),(C12N5/10,  
C12R1:91),  
PC (C12P21/02,C12R1:185),(C12P21/02,C12R1:91),C12N5/00,C12N5/00,  
PC (C12N5/00,C12R1:91)  
CC  
FH Key Location/Qualifiers  
FT CDS (402)..(1331).  
FEATURES  
source  
BASE COUNT 681 a 698 c 669 g 727 t  
ORIGIN  
Query Match 100.0%; Score 931; DB 6; Length 2775;  
Best Local Similarity 100.0%; Pred. No. 1.6e-242;  
Matches 931; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AATGGCTTCCCGAAGATGAGATTGATGATATTTGCTTCTGTTCTGGGGCTCTTTG 60  
Db 401 AATGGCTTCCCGAAGATGAGATTGATGATATTTGCTTCTGTTCTGGGGCTCTTTG 460  
Qy 61 TTGTGATTTACATGTACAGTCAATCTTCAAGACAGTCCCTTTGTTTACAAGAA 120  
Db 461 TTGTGATTTACATGTACAGTCAATCTTCAAGACAGTCCCTTTGTTTACAAGAA 520  
Qy 121 AGACGGGAACCTTCAAGTCCAGATACAGACTGAGCGACACACCTCCCTCTCTCGT 180  
Db 521 AGACGGGAACCTTCAAGTCCAGATACAGACTGAGCGACACACCTCCCTCTCTCGT 580  
Qy 181 COTGCTGTGACCTCATCCCAAAACAGTTGGTGGCGGATGCCATCGCGACAGCGTG 240  
Db 581 COTGCTGTGACCTCATCCCAAAACAGTTGGTGGCGGATGCCATCGCGACAGCGTG 640  
Qy 241 GGGGAAGAGAGATGTTGAAGGAAGAGAGCTGAGACATTTCTTCTCTGGGGACAC 300  
Db 641 GGGGAAGAGAGATGTTGAAGGAAGAGAGCTGAGACATTTCTTCTCTGGGGACAC 700  
Qy 301 CAGCAGTCAGCGGAACCAAGAGGTGACAGGAGGACGAGGACGAGGACGAGGACATTAT 360  
Db 701 CAGCAGTCAGCGGAACCAAGAGGTGACAGGAGGACGAGGACGAGGACGAGGACATTAT 760  
Qy 361 CCAGAAGGATTTCCCTAGAGCTGATTTACATCTGACCTGAACACCATGATGGGCATAGA 420  
Db 761 CCAGAAGGATTTCCCTAGAGCTGATTTACATCTGACCTGAACACCATGATGGGCATAGA 820  
Qy 421 ATGGTCCATCGCTTTTGCTCAGCGCGGTTTGTGATGAAACAGACTCAGACATGTT 480  
Db 821 ATGGTCCATCGCTTTTGCTCAGCGCGGTTTGTGATGAAACAGACTCAGACATGTT 880  
Qy 481 CATCAATGTTGACTATCTGACTGAATGCTTCTGAAGAAAACAGACACACCGAGTTT 540  
Db 881 CATCAATGTTGACTATCTGACTGAATGCTTCTGAAGAAAACAGACACACCGAGTTT 940  
Qy 541 CACTGGCTCTTTGAAACTCAATCAGTTTCCCATCAGGACGCAATTCAGCAAGTGGTTGT 600  
Db 941 CACTGGCTCTTTGAAACTCAATCAGTTTCCCATCAGGACGCAATTCAGCAAGTGGTTGT 1000  
Qy 601 CAGTAAATCTGAATATCCGTGGGACAGGTACCCACCAATTCGTCCCGGACCGGCTAGCT 660  
Db 1001 CAGTAAATCTGAATATCCGTGGGACAGGTACCCACCAATTCGTCCCGGACCGGCTAGCT 1060  
Qy 661 GTTTCTGCGGACGTGGGAGTCAGGTGTACATGTTCTCAAGACGCTCCCATACATPAA 720  
Db 1061 GTTTCTGCGGACGTGGGAGTCAGGTGTACATGTTCTCAAGACGCTCCCATACATPAA 1120  
Qy 721 ACTGAAGACGTGTTTGTGGGGCTCTGCCCTGAAAGCTGAACATCAGATTGGAGAGCT 780  
Db 1121 ACTGAAGACGTGTTTGTGGGGCTCTGCCCTGAAAGCTGAACATCAGATTGGAGAGCT 1180  
Qy 781 CCACCTCCCGAGCGACCTTTTTCAGGGGGCTTACGCTTCTCCGATGCTCTTCAGGAG 840  
Db 1181 CCACCTCCCGAGCGACCTTTTTCAGGGGGCTTACGCTTCTCCGATGCTCTTCAGGAG 1240

Qy 841 GATCGTGGCTGCCACTTTCATCAAGCTCGGACTCTCTTGACACTACTGCGAGGCTCTAGA 900  
Db 1241 GATCGTGGCTGCCACTTTCATCAAGCTCGGACTCTCTTGACACTACTGCGAGGCTCTAGA 1300  
Qy 901 GAATCCCGGGGGGAAGATTGTCGCGCTGTC 931  
Db 1301 GAATCCCGGGGGGAAGATTGTCGCGCTGTC 1331  
RESULT 3  
LOCUS AF145784 933 bp DNA linear PRI 14-NOV-1999  
DEFINITION Homo sapiens beta1,3 galactosyltransferase-v (B3GALT5) gene,  
complete cds.  
ACCESSION AF145784  
VERSION AF145784.1 GI:6409192  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 933)  
AUTHORS Zhou,D., Berger,E.G. and Hennet,T.  
TITLE Molecular cloning of a human UDP-galactose:GlcNAc-beta1,3galNAc  
beta1,3 galactosyltransferase gene encoding an O-linked  
core3-elongation enzyme  
JOURNAL Eur. J. Biochem. 263 (2), 571-576 (1999)  
MEDLINE 99337698  
PUBMED 10406968  
REFERENCE 2 (bases 1 to 933)  
AUTHORS Zhou,D. and Hennet,T.  
TITLE Direct Submission  
JOURNAL Submitted (26-APR-1999) Physiology, University of Zurich,  
Winterthurerstrasse 190, Zurich 8057, Switzerland  
FEATURES  
source  
1..933  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22"  
1..933  
/gene="B3GALT5"  
1..933  
/note="glycosyltransferase"  
/codon\_start=1  
/product="beta1,3 galactosyltransferase-v"  
/protein\_id="AA07880.1"  
/db\_xref="GI:6409193"  
/translation="MAPPKRLMYICLLVLGALCLYFSMYSLNPFKEQSFVYKKGDNF  
LKLPTDCRQTPPLVLLVSSHQLAERMAIRQTKRMVYKQKLTFFLLGTTSS  
AAETKEVDQESQRHGDIIQKDLVYVNLTKTMGIEWHVFPCQAAAFVMTDSDF  
INVDYLTLLKKNRTRFTFLKNEFPFQPFKSWFVSKSEYVPPWDRYPFPCSTG  
YVPSGDVASOVNVKSPVYIKLEDFVGLCLERLNIIRLELHSQPTFFPGGLRESVC  
LFRIVACHFIKPRLLDYQALNSRGDCPPV"  
BASE COUNT 227 a 232 c 241 g 233 t  
ORIGIN  
Query Match 99.9%; Score 930; DB 9; Length 933;  
Best Local Similarity 100.0%; Pred. No. 2.9e-242;  
Matches 930; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2 ATGGCTTCCCGAAGATGAGATTGATGATATTTGCTTCTGTTCTGGGGCTCTTTG 61  
Db 1 ATGGCTTCCCGAAGATGAGATTGATGATATTTGCTTCTGTTCTGGGGCTCTTTG 60  
Qy 62 TTGTATTTAGCATGTACAGTCTAAATCCTTTCAAGAACAGTCTCTTTTACAAGAA 121  
Db 61 TTGTATTTAGCATGTACAGTCTAAATCCTTTCAAGAACAGTCTCTTTTACAAGAA 120  
Qy 122 GACGGGACCTTCTTAAGTCCCGAGATACAGACTCCAGCAGACACCTCCCTCTCTC 181  
Db 121 GACGGGACCTTCTTAAGTCCCGAGATACAGACTCCAGCAGACACCTCCCTCTCTC

Db	121	GACGGGAACCTTCCTTTAAGCTCCACAGATACAGACTGCAGGACGACACCTCCCTTCTCCTGTC	181
QY	182	CTGCTGGTGACCTCATCCACAAACAGTTGGCTGAGCGCATGGCCATCCGGCAGACGCTGG	241
Db	181	CTGCTGGTGACCTCATCCACAAACAGTTGGCTGAGCGCATGGCCATCCGGCAGACGCTGG	240
QY	242	GGGAAGAGAGAGGATGGTGAAGGGAAGACAGCTGAAGACATCTTCTCTCTGCGGACACC	301
Db	241	GGGAAGAGAGAGGATGGTGAAGGGAAGACAGCTGAAGACATCTTCTCTCTGCGGACACC	300
QY	302	AGCAGTCAGCGGGAACGAAAGAGGTGGACGAGAGACGACGACGCGGGACATTATC	361
Db	301	AGCAGTCAGCGGGAACGAAAGAGGTGGACGAGAGACGACGACGCGGGACATTATC	360
QY	362	CAGAAGGATTTCCTAGACGCTATTACAATCTCACCTCGAAGACCATGATGGGCATAGAA	421
Db	361	CAGAAGGATTTCCTAGACGCTATTACAATCTCACCTCGAAGACCATGATGGGCATAGAA	420
QY	422	TGGGTCCATCGCTTTTGTCTCAGCGGGCGTTTGTGATGAAACAGACTCAGACATGTC	481
Db	421	TGGGTCCATCGCTTTTGTCTCAGCGGGCGTTTGTGATGAAACAGACTCAGACATGTC	480
QY	482	ATCAATGTTGACTATCTGACTGAACTGCTCTCGAAGAAAAACAGAACACCCAGGTTTTTC	541
Db	481	ATCAATGTTGACTATCTGACTGAACTGCTCTCGAAGAAAAACAGAACACCCAGGTTTTTC	540
QY	542	ACTGCTCTCTTGAAACTCAATGAGTTTCCCATCAGGACGACATTCAGCAAGTGGTTGTC	601
Db	541	ACTGCTCTCTTGAAACTCAATGAGTTTCCCATCAGGACGACATTCAGCAAGTGGTTGTC	600
QY	602	AGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTCTGCTCCGGCACCGGCTACGTG	661
Db	601	AGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTCTGCTCCGGCACCGGCTACGTG	660
QY	662	TTTTCTGGGACGTGCGGAGTCAGGTGTACAATGCTCTCCAGAGGGTCCCATACATTAAA	721
Db	661	TTTTCTGGGACGTGCGGAGTCAGGTGTACAATGCTCTCCAGAGGGTCCCATACATTAAA	720
QY	722	CTGGAAGACGTGTTCTGGGGCTGCTGCTCGAAGAGCTGAACATCAGATTGGAGGAGCTC	781
Db	721	CTGGAAGACGTGTTCTGGGGCTGCTGCTCGAAGAGCTGAACATCAGATTGGAGGAGCTC	780
QY	782	CATCCCCAGCCGACCTTTTTTCCAGGGGGCTTAGCTTCTCCGTATGCCTCTTCAGGAGG	841
Db	781	CATCCCCAGCCGACCTTTTTTCCAGGGGGCTTAGCTTCTCCGTATGCCTCTTCAGGAGG	840
QY	842	ATCGTGCCCTGCCACTTCATCAGCCTCGGACCTCTCTTGACTACTGCGAGCTCTAGAG	901
Db	841	ATCGTGCCCTGCCACTTCATCAGCCTCGGACCTCTCTTGACTACTGCGAGCTCTAGAG	900
QY	902	AATTCCCGGGGGAAGATTGTCGCCCTGTC	931
Db	901	AATTCCCGGGGGAAGATTGTCGCCCTGTC	930
RESULT	4		
LOCUS	E38420		
DEFINITION	Novel polypeptide.	10562 bp	DNA linear PAT 31-JAN
ACCESSION	E38420		
VERSION	E38420.1	GI:18626994	
KEYWORDS	JP 2000245464-A/2.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 10562)		
JOURNAL	Narimatsu, H., Isshiki, S., Togayauchi, A. and Sasaki, K.		
COMMENT	Novel polypeptide		
	Patent: JP 2000245464-A 2 12-SEP-2000;		
	OS Homo sapiens (human)		
	PN JP 2000245464-A/2		

QY	662	TTTCTCGGAGACGTGGCGAGTCAGGTGTACAATGTCTCCCAAGAGCGTCCCACATATAA	721
Df	8894	TTTTCTCGGGACGTGGCGAGTCAGGTGTACAATGTCTCCCAAGAGCGTCCCACATATAA	8953
QY	722	CTGGAAGACGTGTTCTGTGGGGCTCTGCCCTCGAAAGCGTGAACATCAGATTGGAGGAGCTC	781
Df	8954	CTGGAAGACGTGTTCTGTGGGGCTCTGCCCTCGAAAGCGTGAACATCAGATTGGAGGAGCTC	9013
QY	782	CACTCCAGCCGACCCTTTTTCACGGGGGCTTACGGCTTCTCCGTATGCCCTCTTCAGGAGG	841
Df	9014	CACTCCAGCCGACCCTTTTTCACGGGGGCTTACGGCTTCTCCGTATGCCCTCTTCAGGAGG	9073
QY	842	ATCGTGCCCTGCACCTTCATCAAGCCTCGGACTCTCTTGAGCTACTGGCAGGCTCTAGAG	901
Df	9074	ATCGTGCCCTGCACCTTCATCAAGCCTCGGACTCTCTTGAGCTACTGGCAGGCTCTAGAG	9133
QY	902	AATTCCGGGGGGAAGAATTGTCGCCCTGTC	931
Df	9134	AATTCCGGGGGGAAGAATTGTCGCCCTGTC	9163

RESULT 5  
AF064860 LOCUS  
DEFINITION Homo sapiens chromosome 21 clone PAC 70134 map 21q22.3, complete sequence.  
ACCESSION AF064860  
VERSION AF064860.2 GI:18958624  
KEYWORDS HTG; HTGS\_DRAFT.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 170121)  
Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T., Park,H.-S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.-K., Soeda,E., Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K., Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R., Patterson,D., Reichwald,K., Rump,A., Schillhaber,M.B., Schudy,A., Zimmermann,W., Rosenthal,A., Kudoh,J., Kawasaki,K., Asakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuyama,S., Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G., Hornischer,K., Brandt,P., Schafie,M., Schoen,O., Desario,A., Reichelt,J., Kauer,G., Bloeker,H., Ramser,J., Beck,A., Klages,S., Hennig,S., Riesselmann,L., Dagand,E., Haaf,T., Wehrmeyer,S., Borzym.K., Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and Yaspo,M.Laure.  
The DNA sequence of human chromosome 21 Nature 405 (6784), 311-319 (2000)

TITLE Nature 405 (6784), 311-319 (2000)  
MEDLINE 20289799  
PUBMED 10830953

REFERENCE 2 (bases 1 to 170121)  
Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B., Weber,J., Schattevoy,R., Yaspo.M.-L. and Rosenthal,A.  
Direct Submission

TITLE Submitted (12-MAY-1998) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany  
JOURNAL 3 (bases 1 to 170121)  
Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B., Weber,J., Schattevoy,R., Yaspo.M.-L., Rosenthal,A., Yaspo,M.-L. and Rosenthal,A.

REFERENCE Direct Submission  
JOURNAL Submitted (27-FEB-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany  
REFERENCE 4 (bases 1 to 170121)  
Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B., Weber,J., Schattevoy,R., Yaspo,M.-L. and Rosenthal,A.  
Direct Submission

TITLE Submitted (05-MAR-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany  
JOURNAL On Feb 27, 2002 this sequence version replaced gi:3171153.  
COMMENT Location/Qualifiers



/note="L1MD3"  
/rpt\_family="LINE/L1"  
/rpt\_type=DISPERSED  
complement(8482. .8762)  
/note="AluJo"  
/rpt\_family="SINE/Alu"  
/rpt\_type=DISPERSED  
complement(8865. .8964)  
/note="L2"  
/rpt\_family="LINE/L2"  
/rpt\_type=DISPERSED  
9131. .9424  
/note="AluSq"  
/rpt\_family="SINE/Alu"  
/rpt\_type=DISPERSED  
complement(9723. .9882)  
/note="L2"  
/rpt\_family="LINE/L2"  
/rpt\_type=DISPERSED  
complement(9989. .10077)  
/note="L2"  
/rpt\_family="LINE/L2"  
/rpt\_type=DISPERSED  
complement(10078. .10429)  
/note="THE1B"  
/rpt\_family="LTR/MaLR"  
/rpt\_type=DISPERSED  
complement(10430. .10651)  
/note="L2"  
/rpt\_family="LINE/L2"  
/rpt\_type=DISPERSED  
10980. .11000  
/note="TTTA)n"  
/rpt\_family="Simple\_repeat"  
/rpt\_type=JANDEM  
complement(11002. .11282)  
/note="AluSp"  
/rpt\_family="SINE/Alu"  
/rpt\_type=DISPERSED  
complement(11315. .11556)  
/note="AluJo"  
/rpt\_family="SINE/Alu"  
/rpt\_type=DISPERSED  
11702. .11753  
/note="MIR"  
/rpt\_family="SINE/MIR"  
/rpt\_type=DISPERSED  
12649. .12678  
/gene="SH3BGR"  
/number=5  
complement(13115. .13425)  
/note="AluSx"  
/rpt\_family="SINE/Alu"  
/rpt\_type=DISPERSED  
complement(13598. .13892)  
/note="AluY"  
/rpt\_family="SINE/Alu"  
/rpt\_type=DISPERSED  
14013. .14131  
/note="CA)n"  
/rpt\_family="Simple\_repeat"

Query Match 99.9%; Score 930; DB 9; Length 340000;  
Best Local Similarity 100.0%; Pred. No. 3.3e-242;  
Matches 930; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 ATGGCTTTCCCAAGATGATGATGATATATTGCTTCTGGTTCGGGGCTCTTTGT 61  
Db 163456 ATGGCTTTCCCAAGATGATGATGATATATTGCTTCTGGTTCGGGGCTCTTTGT 163515  
QY 62 TTGATTTTAGCATGTACAGTCTAAATCTTCAAGAACAGACGCTTTGTTTACAAGAAA 121  
Db 163516 TTGATTTTAGCATGTACAGTCTAAATCTTCAAGAACAGACGCTTTGTTTACAAGAAA 163575

QY 122 GACGGGAACCTTCTTAAGCTCCAGATACAGACTGCAGGACAGACACCTCCCTTCCTCGTC 181  
Db 163576 GACGGGAACCTTCTTAAGCTCCAGATACAGACTGCAGGACAGACACCTCCCTTCCTCGTC 163635  
QY 182 CTGCTGGTACCTCATCCACAAACAGTTGGCTGAGCGCATGGCCATCGGCGACAGCTGG 241  
Db 163636 CTGCTGGTACCTCATCCACAAACAGTTGGCTGAGCGCATGGCCATCGGCGACAGCTGG 163695  
QY 242 GGAAGAGAGAGGATGGTGAAGGGAAGACAGCTGAAGACATTTCTCTCTCTGGGACACC 301  
Db 163696 GGAAGAGAGAGGATGGTGAAGGGAAGACAGCTGAAGACATTTCTCTCTCTGGGACACC 163755  
QY 302 AGCAGTGCAGCGGAACAGAAAGAGGTGGACAGAGAGCCAGCGACAGCGGACATTTATC 361  
Db 163756 AGCAGTGCAGCGGAACAGAAAGAGGTGGACAGAGAGCCAGCGACAGCGGACATTTATC 163815  
QY 362 CAGAGAGATTTCTAGACGCTTATTACAATCTGACCTGAAGACCATGATGGCATAGAA 421  
Db 163816 CAGAGAGATTTCTAGACGCTTATTACAATCTGACCTGAAGACCATGATGGCATAGAA 163875  
QY 422 TGGTCCATCGCTTTTGTCTCAGGCGGCTTTGTGATGAAAAACAGACTCAGACATGTTTC 481  
Db 163876 TGGTCCATCGCTTTTGTCTCAGGCGGCTTTGTGATGAAAAACAGACTCAGACATGTTTC 163935  
QY 482 ATCAATGTTGACTATCTGACTGAATGCTTCTGAAGAAAAACAGAACACCAAGTTTTC 541  
Db 163936 ATCAATGTTGACTATCTGACTGAATGCTTCTGAAGAAAAACAGAACACCAAGTTTTC 163995  
QY 542 ACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGAGCCATTCAGAAAGTGGTTTTC 601  
Db 163996 ACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGAGCCATTCAGAAAGTGGTTTTC 164055  
QY 602 AGTAATCTGAATATCGTGGGACAGGTACCCACCATTTCTGCTCGGACCGGCTACGTG 661  
Db 164056 AGTAATCTGAATATCGTGGGACAGGTACCCACCATTTCTGCTCGGACCGGCTACGTG 164115  
QY 662 TTTTCTGGGACGTGGCGAGTCAAGTGTACAAATGCTCCAAGAGGCTCCCATACATTAA 721  
Db 164116 TTTTCTGGGACGTGGCGAGTCAAGTGTACAAATGCTCCAAGAGGCTCCCATACATTAA 164175  
QY 722 CTGGAAGACGTGTTTGTGGGCTCTGCTCGAAAGGCTGAACATCAGATTGGAGGAGCTC 781  
Db 164176 CTGGAAGACGTGTTTGTGGGCTCTGCTCGAAAGGCTGAACATCAGATTGGAGGAGCTC 164235  
QY 782 CACTCCAGCGGACCTTTTTTCCAGGGGCTTAGCTTCTCGGTATGCTCTTCAGGAG 841  
Db 164236 CACTCCAGCGGACCTTTTTTCCAGGGGCTTAGCTTCTCGGTATGCTCTTCAGGAG 164295  
QY 842 ATCGTGGCTGCCACTTTCATCAAGCCTCGGACTCTCTTGGACTACTGGCAGGCTCTAGAG 901  
Db 164296 ATCGTGGCTGCCACTTTCATCAAGCCTCGGACTCTCTTGGACTACTGGCAGGCTCTAGAG 164355  
QY 902 AATTCGGGGGGAAGATTGTCCGCTGTC 931  
Db 164356 AATTCGGGGGGAAGATTGTCCGCTGTC 164385

RESULT 7  
HSA6078  
LOCUS HSA6078 933 bp DNA linear PRI 11-MAY-2000  
DEFINITION Homo sapiens beta3gal-T6 gene.  
ACCESSION AJ006078  
VERSION AJ006078.1 GI:7799922  
KEYWORDS beta-1,3-galactosyltransferase; beta3gal-T6 gene.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 933)  
AUTHORS Amado, M., Carneiro, F. and Clausen, H.  
TITLE Cloning and expression of two beta-1,3-galactosyltransferases:  
beta3gal-T5 and beta3gal-T6



JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 933)  
AUTHORS Amado, M.  
TITLE Direct Submission  
JOURNAL Submitted (11-MAY-1998) Amado M., Department of Oral Diagnostics,  
Royal Dental School, Norre Alle 20, 2200 Copenhagen, DENMARK

FEATURES  
source 1..933 Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
gene 1..933  
/gene="beta3gal-T6"  
CDS 1..933  
/gene="beta3gal-T6"  
/codon\_start=1  
/product="beta-1,3-galactosyltransferase"  
/protein\_id="CAB91547.1"  
/db\_xref="GI:7799923"  
/translation="MAPPKRLMYICLLVLGALCLYFSMNSLNPKEQSFVYKKDGNF  
LKLPDTCRQTPFLLVTSKQLAERMAIROTGWKERTVKGKOLKTFELLCGTSS  
AAETKEVDQESQHGDIQKDFLVYNLTKTMGIEWHFRFCQAARFVMTDSMF  
INVDYLTLLKLNRTTRFTGFLKLNFFPIQPFKSWFVSKSEYPMWRYPFPCSGTG  
YVSGDVASQVYNVSKSPYIKLEDVFVGLCLERLNIRLEELHSQPTFFPGGLRFVSC  
LFRRIVACHFIKPRITLLDYWALENSRGEDCPV"

BASE COUNT 229 a 234 c 240 g 229 t 1 others  
ORIGIN

Query Match 98.9%; Score 921; DB 9; Length 933;  
Best Local Similarity 99.4%; Pred. No. 8.3e-240;  
Matches 924; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 ATGCTTTCCGGAAGATGATGATATATTTGCTTCTGCTTCTGGGGCTCTTTGT 61  
DB 1 ATGCTTTCCGGAAGATGATGATATATCTGCTTCTGCTTCTGGGGCTCTTTGT 60  
QY 62 TTGATTTTAGCATGTACAGTCTAAATCCTTTCAAGAAGACAGCCCTTTGTTACAGAAA 121  
DB 61 TTGATTTAGCATGAACAGTCTAAATCCTTTCAAGAAGACAGCTTTGTTACAGAAA 120  
QY 122 GACGGAACTTCCTTAAGCTCCCATACAGACTGACGAGCAGACACCTCCCTTCTCTGTC 181  
DB 121 GACGGAACTTCCTTAAGCTCCCATACAGACTGACGAGCAGACACCTCCCTTCTCTGTC 180  
QY 182 CTGCTGGTGACCTCATCCACAAACAGTTGGCTGACGGATGCCATCCGCGACAGCTGG 241  
DB 181 CTGCTGGTGACCTCATCCACAAACAGTTGGCTGACGGATGCCATCCGCGACAGCTGG 240  
QY 242 GGGAAAGAGAGATGGTGAAGGAAAGCAGCTGAAGACATTTCTCTCTGGGGACACC 301  
DB 241 GGGAAAGAGAGAGCTGAAGGAAAGCAGCTGAAGACATTTCTCTCTGGGGACACC 300  
QY 302 ACCAGTGACGGGAACGAAAGAGTGGACAGGAGCCAGGACGACGAGGACATTATC 361  
DB 301 ACCAGTGACGGGAACGAAAGAGTGGACAGGAGCCAGGACGACGAGGACATTATC 360  
QY 362 CAGAGGATTTCTAGACGTCTATTACAATCTGACCTGAAGACCATGATGGGCATAGAA 421  
DB 361 CAGAGGATTTCTAGACGTCTATTACAATCTGACCTGAAGACCATGATGGGCATAGAA 420  
QY 422 TGGGTCCATCGCTTTTGCTCTCAGCGGGCGTTTGATGATGAAAACAGACTCAGACATGTC 481  
DB 421 TGGGTCCATCGCTTTTGCTCTCAGCGGGCGTTTGATGATGAAAACAGACTCAGACATGTC 480  
QY 482 ATCAATGTTGACTATCTGACTGACTGCTTCTGAAAGAAACAGACAAACAGGTTTTC 541  
DB 481 ATCAATGTTGACTATCTGACTGACTGCTTCTGAAAGAAACAGACAAACAGGTTTTC 540  
QY 542 ACTGGCTTCTGAAACTCAATGAGTTTCCCATCAGCAGCAGCATTCAGCAAGTGTTC 601  
DB 541 ACTGGCTTCTGAAACTCAATGAGTTTCCCATCAGCAGCAGCATTCAGCAAGTGTTC 600  
QY 602 AGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGCGACCGGTACGTG 661

601 AGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGCGACCGGTACGTG 660  
QY 662 TTTTCTGGGACGTGGGACGTAGGTGTAATGTTCTCCAGAGCGTCCCATACATTAAA 721  
DB 661 TTTTCTGGGACGTGGGACGTAGGTGTAATGTTCTCCAGAGCGTCCCATACATTAAA 720  
QY 722 CTGGAAGAGCTGTTTGTGGGGCTCTGCTCGAAGGCTGAACATCAGATTGGAGAGCTC 781  
DB 721 CTGGAAGAGCTGTTTGTGGGGCTCTGCTCGAAGGCTGAACATCAGATTGGAGAGCTC 780  
QY 782 CACTCCAGCCGACCTTTTTCAGGGGGCTTACGCTTCTCCGCTATGCTCTTCAGGAGG 841  
DB 781 CACTCCAGCCGACCTTTTTCAGGGGGCTTACGCTTCTCCGCTATGCTCTTCAGGAGG 840  
QY 842 ATCTGGCTGCTGCTCATCATCAAGCCTCGGACTCTCTTGGACTACTGGAGGCTCTAGAG 901  
DB 841 ATCTGGCTGCTGCTCATCATCAAGCCTCGGACTCTCTTGGACTACTGGAGGCTCTAGAG 900  
QY 902 AATTCCTGGGGGGAAGATTGTCGCCCTGTC 931  
DB 901 AATTCCTGGGGGGAAGATTGTCGCCCTGTC 930

RESULT 8  
AB041416 1576 bp DNA linear PRI 17-OCT-2000  
LOCUS Homo sapiens beta1,3-GalT 5 gene for UDP-Gal:GlcNAc  
DEFINITION beta1,3-galactosyltransferase 5, partial cds.  
ACCESSION AB041416  
VERSION AB041416.1 GI:7593026  
KEYWORDS  
SOURCE Homo sapiens (isolate:#056) DNA.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1576)  
AUTHORS Liu, Y. and Saitou, N.  
TITLE Silver Project  
JOURNAL Published Only in DataBase (2000)  
REFERENCE 2 (bases 1 to 1576)  
AUTHORS Liu, Y. and Saitou, N.  
TITLE Direct Submission  
JOURNAL Submitted (11-APR-2000) Naruya Saitou, National Institute of  
Genetics, Laboratory of Evolutionary Genetics, 1111 Yata, Mishima,  
Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,  
URL:http://sayer.lab.nig.ac.jp/~silver/, Tel:81-559-81-6790,  
Fax:81-559-81-6789)  
FEATURES Location/Qualifiers  
1..1576  
/organism="Homo sapiens"  
/isolate="#056"  
/db\_xref="taxon:9606"  
/note="human sequence used for primer design based on Acc#  
AB020337"  
exon <1..27  
/number=3  
intron 28..674  
/number=3  
gene 675..1576  
/gene="beta1,3-GalT 5"  
CDS 675..>1576  
/gene="beta1,3-GalT 5"  
/codon\_start=1  
/product="UDP-Gal:GlcNAc beta1,3-galactosyltransferase 5"  
/protein\_id="BAA94501.1"  
/db\_xref="GI:7593027"  
/translation="MAPPKRLMYICLLVLGALCLYFSMNSLNPKEQSFVYKKDGNF  
LKLPDTCRQTPFLLVTSKQLAERMAIROTGWKERTVKGKOLKTFELLCGTSS  
AAETKEVDQESQHGDIQKDFLVYNLTKTMGIEWHFRFCQAARFVMTDSMF  
INVDYLTLLKLNRTTRFTGFLKLNFFPIQPFKSWFVSKSEYPMWRYPFPCSGTG  
YVSGDVASQVYNVSKSPYIKLEDVFVGLCLERLNIRLEELHSQPTFFPGGLRFVSC  
LFRRIVACHFIKPRITLLDYWALENSRGEDCPV"  
exon 675..>1576

/gene="betal,3-Galt 5" /number=4														
BASE COUNT	415 a	373 c	387 g	401 t										
ORIGIN														
Query Match	96.5%; Score 898.8; DB 9; Length 1576;													
Best Local Similarity	99.8%; Pred. No. 9.4e-234;													
Matches	900; Conservative	0; Mismatches	2; Indels	0; Gaps	0;									
QY	2	ATGGCTTTCCCGAAGATGAGATTGATGATATATTGGCTTCTGGTTCTGGGGGCTCTTTGT	61											
Db	675	ATGGCTTTCCCGAAGATGAGATTGATGATATATTGGCTTCTGGTTCTGGGGGCTCTTTGT	734											
QY	62	TTGTATTTAGCATGTACAGTCTAAATCCTTTCAAGAACAGTCCCTTTGTTTACAGAAA	121											
Db	735	TTGTATTTAGCATGTACAGTCTAAATCCTTTCAAGAACAGTCCCTTTGTTTACAGAAA	794											
QY	122	GACGGGAACCTTCCCTTAAGCTCCCATACAGACTGCAGGACAGACACTCCCTTCCTCGTC	181											
Db	795	GACGGGAACCTTCCCTTAAGCTCCCATACAGACTGCAGGACAGACACTCCCTTCCTCGTC	854											
QY	182	CTGCTGGTGACCTCATCCACAAACAGTTGGCTGAGCGCATGCCATCCGGCAGACGTGG	241											
Db	855	CTGCTGGTGACCTCATCCACAAACAGTTGGCTGAGCGCATGCCATCCGGCAGACGTGG	914											
QY	242	GGGAAAGAGAGGATGGTGAAGGAAAGCAGCTGAAGACATTCTTCCTCGGGGACACC	301											
Db	915	GGGAAAGAGAGGATGGTGAAGGAAAGCAGCTGAAGACATTCTTCCTCGGGGACACC	974											
QY	302	AGCAGTGCAGCGGAACGAAGAGGTGGACAGGAGGACGCGACACGGGGACATTATC	361											
Db	975	AGCAGTGCAGCGGAACGAAGAGGTGGACAGGAGGACGCGACACGGGGACATTATC	1034											
QY	362	CAGAAGGATTTCTAGACGTCTATTACATCTGACCTCAGACCATGATGGGCATAGAA	421											
Db	1035	CAGAAGGATTTCTAGACGTCTATTACATCTGACCTCAGACCATGATGGGCATAGAA	1094											
QY	422	TGGGTCCATCGCTTTTGCTCCTCAGCGGGCTTGTGATGAAACAGACTCAGACATGTC	481											
Db	1095	TGGGTCCATCGCTTTTGCTCCTCAGCGGGCTTGTGATGAAACAGACTCAGACATGTC	1154											
QY	482	ATCAATGTTGACTATCTGACTGAATGCTTCTGAAGAAAACAGAACACAGGTTTTTC	541											
Db	1155	ATCAATGTTGACTATCTGACTGAATGCTTCTGAAGAAAACAGAACACAGGTTTTTC	1214											
QY	542	ACTGGCTTCTTGAAGCTCAATGAGTTTCCATCAGGACGCCATTACAGCAAGTGGTTGTC	601											
Db	1215	ACTGGCTTCTTGAAGCTCAATGAGTTTCCATCAGGACGCCATTACAGCAAGTGGTTGTC	1274											
QY	602	AGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGGCACCGCTACGTG	661											
Db	1275	AGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGGCACCGCTACGTG	1334											
QY	662	TTTCTGGCGAGCTGGCGAGTCAAGTGTACAATGTCTCCAAGAGCGTCCCATACATTAA	721											
Db	1335	TTTCTGGCGAGCTGGCGAGTCAAGTGTACAATGTCTCCAAGAGCGTCCCATACATTAA	1394											
QY	722	CTGGAAGACGTGTTTGTGGGGCTCTGCCCTCGAAAGCGCTGAACATCAGATTGGAGAGCTC	781											
Db	1395	CTGGAAGACGTGTTTGTGGGGCTCTGCCCTCGAAAGCGCTGAACATCAGATTGGAGAGCTC	1454											
QY	782	CACATCCAGCGACCTTTTTCAGGGGGCTTACGGTTCTCCGTATGCCCTTCAGGAGG	841											
Db	1455	CACATCCAGCGACCTTTTTCAGGGGGCTTACGGTTCTCCGTATGCCCTTCAGGAGG	1514											
QY	842	ATCGTGGCTGCCACTTCATCAAGCCTCGGACTCTCTTGGACTACTGGCAGGCTCTAGAG	901											
Db	1515	ATCGTGGCTGCCACTTCATCAAGCCTCGGACTCTCTTGGACTACTGGCAGGCTCTAGAG	1574											
QY	902	AA 903												
Db	1575	AA 1576												

/gene="betal,3-Galt 5" /isolate="bonobo-05" /db_xref="taxon:9597" /note="human sequence used for primer design based on Acc#														
AB020337"														
<1..20														
/number=3														
21..667														
/number=3														
668..1570														
/gene="betal,3-Galt 5"														
668..>1570														
/gene="betal,3-Galt 5"														
/codon_start=1														
/product="UDP-Gal:GlcNAc betal,3-galactosyltransferase 5"														
/protein_id="BAA94500.1"														
/db_xref="GI:7593025"														
/translation="MAFPRMLMYVCLLVGLALCLYPSMYSINLFKEQSFYKKDGNF														
LKLPDTCRQTPFLVLLVTSKQLAERMAIRQWERTVKGKQLATFLLGTSS														
AAETKEVDQESORHGDIOKDFLDGYNLTLMKMMIEWHRFCQAAAVMKRSDMF														
INVDYLTLLKKNRTTREFGLKNEFPIQSKFWKSVKSPMDRYPPFCSTG														
YVSGDVASQVYNVSESPVYKLEDFVGLCLERLNIRLELHSLQTPFPGLRFSVC														
RFRRIVACHEIKPRLLDYWQALEN"														
668..>1570														
/gene="betal,3-Galt 5"														
/number=4														
BASE COUNT	408 a	372 c	392 g	398 t										
ORIGIN														
Query Match	94.8%; Score 882.2; DB 9; Length 1570;													
Best Local Similarity	98.6%; Pred. No. 3.1e-229;													
Matches	890; Conservative	0; Mismatches	13; Indels	0; Gaps	0;									
QY	2	ATGGCTTTCCCGAAGATGAGATTGATGATATATTGGCTTCTGGTTCTGGGGGCTCTTTGT	61											
Db	668	ATGGCTTTCCCGAAGATGAGATTGATGATATATTGGCTTCTGGTTCTGGGGGCTCTTTGT	727											
QY	62	TTGTATTTAGCATGTACAGTCTAAATCCTTTCAAGAACAGTCTCTTTGTTTACAGAAA	121											
Db	728	TTGTATTTAGCATGTACAGTCTGAAATCTTTTCAAGAACAGTCTCTTTGTTTACAGAAA	787											
QY	122	GACGGGAACCTTCCCTTAAGCTCCCATACAGACTGCAGGACAGACACTCCCTTCCTCGTC	181											
Db	788	GATGGGAACCTTCCCTTAAGCTCCCATACAGACTGCAGGACAGACACTCCCTTCCTCGTC	847											
QY	182	CTGCTGGTGACCTCATCCACAAACAGTTGGCTGAGCGCATGGCCATCCGGCAGACGTGG	241											



Qy 542 ACTGCGCTTCTTGAAGTCAATGAGTTCCCATCAGGAGCCATTACGAAAGTGGTTGTC 601  
Db 1215 ACTGCGCTTCTTGAAGTCAATGAGTTCCCATCAGGAGCCATTACGAAAGTGGTTGTC 1274  
Qy 602 AGTAATCTGAATACCGTGGACAGGTACCCACCATCTGCTCCGGCACCGGTACGTCG 661  
Db 1275 AGTAATCTGAATACCGTGGACAGGTACCCACCATCTGCTCCGGCACCGGTACGTCG 1334  
Qy 662 TTTTCTGGCGAGCTGGCGAGTCAAGTGTACAATGTCTCCAAGAGCGTCCCATACATTAAA 721  
Db 1335 TTTTCTGGCGAGCTGGCGAGTCAAGTGTACAATGTCTCCAAGAGCGTCCCATACATTAAA 1394  
Qy 722 CTGGAAGAGCTGTTGTGGGGCTCTGCCCTCGAAAGGCTGAACATCAGATTGGAGAGCTC 781  
Db 1395 CTGGAAGAGCTGTTGTGGGGCTCTGCCCTCGAAAGGCTGAACATCAGATTGGAGAGCTC 1454  
Qy 782 CACTCCCGACCGACCTTTTTCAGGGGCTTACGCTTCTCCGATCTCCGTCCTTCAGGAGG 841  
Db 1455 CACTCCCGACCGACCTTTTTCAGGGGCTTACGCTTCTCCGATCTCCGTCCTTCAGGAGG 1514  
Qy 842 ATCTGGCGCTGCACCTTCATCAAGCTCGGACTCTCTTGGACTACTGGCAGG 893  
Db 1515 ATCTGGCGCTGCACCTTCATCAAGCTCGGACTCTC-TGGACTACTGGCAGG 1565  
  
RESULT 11  
AB041412  
LOCUS AB041412 1360 bp DNA linear PRI 17-OCT-2000  
DEFINITION Gorilla gorilla betal,3-Galt 5' gene for UDP-Gal:GlcNAc  
betal,3-galactosyltransferase 5, partial cds.  
ACCESSION AB041412  
VERSION AB041412.1 GI:7593018  
KEYWORDS  
SOURCE Gorilla gorilla (isolate:#085) DNA.  
ORGANISM Gorilla gorilla  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Gorilla.  
REFERENCE 1 (bases 1 to 1360)  
AUTHORS Liu,Y. and Saitou,N.  
JOURNAL Silver Project  
TITLE Published Only in DataBase (2000)  
REFERENCE 2 (bases 1 to 1360)  
AUTHORS Liu,Y. and Saitou,N.  
JOURNAL Direct Submission  
TITLE Submitted (11-APR-2000) Naruya Saitou, National Institute of  
Genetics, Laboratory of Evolutionary Genetics, 1111 Yata, Mishima,  
Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,  
URL:http://sayer.lab.nig.ac.jp/~silver/, Tel:81-559-81-6790,  
Fax:81-559-81-6789)  
  
FEATURES  
source  
1..1360  
/organism="Gorilla gorilla"  
/isolate="#085"  
/db\_xref="taxon:9593"  
/note="human sequence used for primer design based on Acc#  
AB020337"  
intron  
<1..466  
/number=3  
gene  
467..1360  
/gene="betal,3-Galt 5"  
CDS  
467..>1360  
/gene="betal,3-Galt 5"  
/codon\_start=1  
/product="UDP-Gal:GlcNAc betal,3-galactosyltransferase 5"  
/protein\_id="BAA94497.1"  
/db\_xref="GI:7593019"  
/translation="MAFPAKRLMYICLLVLGALCLIFSYSNLNPFKQSFYKKDVNF  
LKLPDTRQTPFPFLVLLVTSKQLAERMAIRQTWKGKRVKQKLTFFLLGTTSS  
AAETKEVQDSRRHGDIIQKDFLDVYVNLTKTMGIEWHVFQPOAFAVNMKTDSDMF  
INVDYLTLLKNNRTREFTFLKLNERPFPFSKWFVSKSEYPNDRYPFCSGTG  
YVFGSDVASOVNVSESPYIKLEDVFGVLCLERLNIRLEELHLSQPTFFPGGLRFSVC  
RFRRIVACHFIKPRILLDYQQA"

exon 467..>1360  
/gene="betal,3-Galt 5"  
BASE COUNT 334 a 331 c 341 g 351 t 3 others  
ORIGIN  
Query Match 93.78; Score 872; DB 9; Length 1360;  
Best Local Similarity 98.2%; Pred. No. 1.9e-226;  
Matches 878; Conservative 3; Mismatches 13; Indels 0; Gaps 0;  
Qy 2 ATGGCTTTCGCCGAAGATGAGATTGATGTATATTTGCCCTTCTGGTCTGGGGGCTCTTTGT 61  
Db 467 ATGGCTTTCGCCGAAGATGAGATTGATGTATATTTGCCCTTCTGGTCTGGGGGCTCTTTGT 526  
Qy 62 TTGTATTTTACGATGTACAGTCTAAATCCTTTTCAAGAACAGTCCCTTTGTTTACAGAAA 121  
Db 527 TTGTATTTTACGATGTACAGTCTAAATCCTTTTCAAGAACAGTCCCTTTGTTTACAGAAA 586  
Qy 122 GAGGGGAACCTTCCCTTAAGCTCCAGATACAGACTGCAGGCAGACACTCCCTTCCCTGTC 181  
Db 587 GAGCTGAACCTTCCCTTAAGCTCCAGATACAGACTGCAGGCAGACACTCCCTTCCCTGTC 646  
Qy 182 CTGCTGTGACCTCATCCCAACACAGTTGGCTGAGCGCATGGCCATCGGAGACGTGG 241  
Db 647 CTGCTGTGACCTCATCCCAACACAGTTGGCTGAGCGCATGGCCATCGGAGACGTGG 706  
Qy 242 GGGAAAGAGAGGATGGTGAAGGAAAGCAGCTGAAGACATTTCTTCTCTGGGACACC 301  
Db 707 GGGAAAGAGAGGATGGTGAAGGAAAGCAGCTGAAGACATTTCTTCTCTGGGACACC 766  
Qy 302 AGCAGTCCAGCGGAAACGAAAGAGGTGGACCCAGGAGAGCCAGCAGACCGGGGACATTATC 361  
Db 767 AGCAGTCCAGCGGAAACGAAAGAGGTGGACCCAGGAGAGCGCGGACACGGGACATTATC 826  
Qy 362 CAGAAGATTTCTTAGAGCTCTATTACATCTGACCTGAGACCATGATGGGCATAGAA 421  
Db 827 CAGAAGATTTCTTAGAGCTCTATTACATCTGACCTGAGACCATGATGGGCATAGAA 886  
Qy 422 TGGTCCATCGCTTTTGTGCTCAGCGCGGTTTGTGATGAAGAACACACATCAGATGTC 481  
Db 887 TGGTCCATCGCTTTTGTGCTCAGCGCGGTTTGTGATGAAGAACACACATCAGATGTC 946  
Qy 482 ATCAATGTTGACTATCTGACTGAACCTGCTCTGAAGAAAACAGAACACAGCTTTTTC 541  
Db 947 ATCAATGTTGACTATCTGACTGAACCTGCTCTGAAGAAAACAGAACACAGCTTTTTC 1006  
Qy 542 ACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGAGCCCATTCAGCAAGTGGTTTGTGTC 601  
Db 1007 ACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGAGCCCATTCAGTAAGTGGTTTGTGTC 1066  
Qy 602 AGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCCGACCCGGCTAGCTG 661  
Db 1067 AGTAAATCTGAATATCCCATGGGACAGGTACCCGCGCATTTCTGCTCCGCGCACCGGTAGCTG 1126  
Qy 662 TTTTCTGGCGAGCTGGCGAGTCAAGTGTACATGCTCTCAAGAGCGTCCCATACATTAAA 721  
Db 1127 TTTTCCGCGAGCTGGCGAGTCAAGTGTACAAATGTCTCCGAGAGCTCCCATACATTAAA 1186  
Qy 722 CTGGAAGAGCTGTTTGTGGGCTCTGCCTCGAAAGGCTGAACATCAGATTGGAGGAGCTC 781  
Db 1187 CTGGAAGAGCTGTTTGTGGGCTCTGCCTCGAAAGGCTGAACATCAGATTGGAGGAGCTC 1246  
Qy 782 CACTCCCGACCGACCTTTTTCAGGGGCTTACGCTTCTCCGATCTCCGTCCTTCAGGAGG 841  
Db 1247 CACTCCCGACCGACSTTTTTCAGGGGCTTACGCTTCTCCTGATGCGCGCTTCAGGAGG 1306  
Qy 842 ATGCTGGCTGCCACTTTCATCAAGCTCGGACTCTCTTGGACTACTGGCAGGCT 895  
Db 1307 ATGCTGGCTGCCACTTTCATCAAGCTCGGACTCTCTTGGACTACTGGCAGGCT 1360  
  
RESULT 12  
AB0411414

LOCUS AB041414 1566 bp DNA linear PRI 13-APR-2000  
DEFINITION Pan troglodytes betal.3-Galt 5 gene for UDP-Gal:GlcNAC  
betal.3-galactosyltransferase 5, partial cds.  
ACCESSION AB041414  
VERSION AB041414.1 GI:7593022  
KEYWORDS  
SOURCE Pan troglodytes (isolate:chimp-202) DNA.  
ORGANISM  
REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
TITLE 1 (bases 1 to 1566)  
JOURNAL Liu, Y. and Saitou, N.  
PUBLISHED Only in DataBase. (2000)  
REFERENCE 2 (bases 1 to 1566)  
AUTHORS Liu, Y. and Saitou, N.  
TITLE Direct Submission  
JOURNAL Submitted (11-APR-2000) Naruya Saitou, National Institute of  
Genetics, Laboratory of Evolutionary Genetics; 1111 Yata, Mishima,  
Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,  
URL: http://sayer.lab.nig.ac.jp/~silver/. Tel:81-559-81-6790,  
Fax:81-559-81-6789)  
FEATURES  
Source Location/Qualifiers  
1. .1566  
/organism="Pan troglodytes"  
/isolate="chimp-202"  
/db\_xref="taxon:9598"  
/note="human sequence used for primer design based on Acc#  
AB020337"  
exon <1..28  
intron /number=3  
29..675  
gene /number=3  
675..1566  
CDS /gene="betal.3-Galt 5"  
675..>1566  
/gene="betal.3-Galt 5"  
/codon\_start=1  
/product="UDP-Gal:GlcNAC betal.3-galactosyltransferase 5"  
/protein\_id="BAA94499.1"  
/db\_xref="GI:7593023"  
/translation="MAPPKRLMYVLLVLGALCVYFSMYSNLNFKESQSFVYKKGNF  
LKLPDIDCRTPPLFLVLLVSSHRLAERMAIOTWGERIVTKGOLKTFLLGTSS  
AAETKEVDQSRHGDIQKDFLVYNNLTLMGIEWHFRCPQAFVMTDSMF  
INVDYLTLLKRNTRFFTKLNEFPFIQRFKFSKSEYPMWRYPPFCSTG  
YVPSGDVASOVNVSEVPYKLEDFVGLERLINIRLEELHSQTPFPFGILRSVC  
RFRIVACHIKERTLLDYQ"  
exon 675..>1566  
/gene="betal.3-Galt 5"  
/number=4  
BASE COUNT 404 a 374 c 394 g 393 t 1 others  
ORIGIN  
Query Match 93.4%; Score 869.2; DB 9; Length 1566;  
Best Local Similarity 98.3%; Pred. No. 1.le-225;  
Matches 877; Conservative 1; Mismatches 14; Indels 0; Gaps 0;  
QY 2 ATGGCTTCCCGAGATGAGATTGATGTATATTTGCCCTCTCGTCTGGGGGCTCTTTGT 61  
DB 675 ATGGCTTCCCGAGATGAGATTGATGTATTTTCCCTCTCGTCTGGGGGCTCTTTGT 734  
QY 62 TTGTATTTTACATGTCAGTCTAAATCTTTTCAAGAACAGTCTTGTTCACAGAA 121  
DB 735 GTGTATTTTACATGTCAGTCTGAATCTTTTCAAGAACAGTCTTGTTCACAGAA 794  
QY 122 GACGGGAACCTTCCCTTAAGCTCCCGATACAGACTCGAGGACACACCTCCCTTCCTCGTC 181  
DB 795 GACGGGAACCTTCCCTTAAGCTCCCGATACAGACTCGAGGACACACCTCCCTTCCTCGTC 854  
QY 182 CTGCTGTGACCTCATCCCAACAACAGTTGGCTGAGCGCATGGCCATCCGGCAGAGCTGG 241  
DB 855 CTGCTGTGACCTCATCCCAACAACAGTTGGCTGAGCGCATGGCCATCCGGCAGAGCTGG 914

QY 242 GGGAAAGAGAGGATGGTGAAGGGAAAGAGCTGAAGACATTTCTTCTCTGGGGACACC 301  
DB 915 GGGAAAGAGAGGAGCGGTGAAGGGAAAGAGCTGAAGACATTTCTTCTCTGGGGACACC 974  
QY 302 AGCAGTGACGGGAAACGAAAGAGGTGGACCGAGAGCCACGACGAGGACATTATC 361  
DB 975 AGCAGTGACGGGAAACGAAAGAGGTGGACCGAGAGCCACGACGAGGACATTATC 1034  
QY 362 CAGAAGGATTTCTTAGACCTCTATTACATCTCACCTGAAGACCATGATGGGATACAA 421  
DB 1035 CAGAAGGATTTCTTAGACCTCTATTACATCTCACCTGAAGACCATGATGGGATACAA 1094  
QY 422 TGGGTCCATCGCTTTTGTCTCAGCGGGGTTTGTGATGAAAAACAGACTCAGACATGTC 481  
DB 1095 TGGGTCCATCGCTTTTGTCTCAGCGGGGTTTGTGATGAAAAACAGACTCAGACATGTC 1154  
QY 482 ATCAATGTTGACTATCTGACTGAATCTTCTCGAAGAAAAACAGACAGTCTTTTC 541  
DB 1155 ATCAATGTTGACTATCTGACTGAATCTTCTCGAAGAAAAACAGACAGTCTTTTC 1214  
QY 542 ACTGCTCTCTTCAAACTCAATGAGTTTCCCATCAGGAGCCATTTCAGCAAGTGGTTGTC 601  
DB 1215 ACTGCTCTCTTGAATCAATGAGTTTCCCATCAGGAGCCATTTCAGTAACTGAGTTGTC 1274  
QY 602 AGTAATCTGAATATCCGTGGACAGGTACCCACCATTTCTGCTCCGACCGGCTACGTG 661  
DB 1275 AGTAATCTGAATATCCGTGGACAGGTACCCGCTTCTGCTCCGACCGGCTACGTG 1334  
QY 662 TTTCTGGGAGCTGGCGAGTCAGGTGTACATGTCTCCAGAGCGTCCCATACATTA 721  
DB 1335 TTTCTGGGAGCTGGCGAGTCAGGTGTACATGTCTCCAGAGCGTCCCATACATTA 1394  
QY 722 CTGGAAGAGCTGTTGTGGGGCTCTGCTCGAAGGCTGAACATCAGATCGAGGAGCTC 781  
DB 1395 CTGGAAGAGCTGTTGTGGGGCTCTGCTCGAAGGCTGAACATCAGATCGAGGAGCTC 1454  
QY 782 CACTCCCGAGCCGACCTTTTTCAGGGGGCTTACGGTTTCTCCGTATGCTCTTCAGGAGG 841  
DB 1455 CACTCCCGAGCCGACCTTTTTCAGGGGGCTTACGGTTTCTCCGTATGCTCTTCAGGAGG 1514  
QY 842 ATCGTGGGCTGCCACTTCATCAAGCCTCGGACTCTCTTGGACTACTGGCAGG 893  
DB 1515 ATCGTGGGCTGCCACTTCATCAAGCCTCGGACTCTCTTGGACTACTGGCAGG 1566  
RESULT 13  
LOCUS AB041417 1579 bp DNA linear PRI 13-APR-2000  
DEFINITION Pongo pygmaeus betal.3-Galt 5 gene for UDP-Gal:GlcNAC  
betal.3-galactosyltransferase 5, partial cds.  
ACCESSION AB041417  
VERSION AB041417.1 GI:7593028  
KEYWORDS  
SOURCE Pongo pygmaeus (isolate:oran-Pol7) DNA.  
ORGANISM  
REFERENCE Pongo pygmaeus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pongo.  
AUTHORS 1 (bases 1 to 1579)  
TITLE Silver Project  
JOURNAL Published Only in DataBase (2000)  
REFERENCE 2 (bases 1 to 1579)  
AUTHORS Liu, Y. and Saitou, N.  
TITLE Direct Submission  
JOURNAL Submitted (11-APR-2000) Naruya Saitou, National Institute of  
Genetics, Laboratory of Evolutionary Genetics; 1111 Yata, Mishima,  
Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,  
URL: http://sayer.lab.nig.ac.jp/~silver/. Tel:81-559-81-6790,  
Fax:81-559-81-6789)  
FEATURES  
Source Location/Qualifiers  
1. .1579  
/organism="Pongo pygmaeus"  
/isolate="oran-Pol7"

/db\_xref="taxon:9600"  
/note="human sequence used for primer design based on Acc# AB020337"  
exon <1..26  
intron /number=3  
27..673  
/number=3  
674..1579  
/gene="beta1,3-GalT 5"  
674..1579  
/gene="beta1,3-GalT 5"  
/codon\_start=1  
/product="UDP-Gal:GlcNAc beta1,3-galactosyltransferase 5"  
/protein\_id="BAA94502.1"  
/db\_xref="GI:7593029"  
/translation="MAFKRMRIYICLLVLGALCLYFSMNLNPFKEOSVYKKEDGN FLKLPDIDCROTPPLVLLVTSQKQLAERMAIROVMGKERTVKGQKLTFFLLGTTSS SAABKEVDQSSQRHDIQKDFLDVYNNLTMTMGMEWHRFCQAAVFMKTDSDM FINVDYJELLKNNKTRTFEFTGLKLNLPQPFKSWFVSEYPMWDRYPFPCSGT AVFSGDVASQVYNSVYPYIKLEDVFGVGLCLERLNIRLELHSHOFTFPFEGLRFSV CRFRIVACHIKPRTLLDYMOALEN"  
674..1579  
exon /gene="beta1,3-GalT 5"  
/number=4  
421 a 371 c 383 g 404 t  
BASE COUNT 421 a 371 c 383 g 404 t  
ORIGIN  
Query Match 91.3%; Score 850; DB 9; Length 1579;  
Best Local Similarity 96.9%; Pred. No. 1.9e-220;  
Matches 878; Conservative 0; Mismatches 25; Indels 3; Gaps 1;  
QY 2 ATGGCTTTCCCGAAGATGAGATTGATGTATATTGGCTTCTGGTCTGGGGCTCTTTGT 61  
DB 674 ATGGCTTTCCCGAAGATGAGATTGATGTATATTGGCTTCTGGTCTGGGGCTCTTTGT 733  
QY 62 TTGATTTTAGCATGTACAGTCTAAATCCTTTCAAGAACAGACGCTTTGTTTACAGAAA 121  
DB 134 TTGATTTTAGCATGTACAATCTGAATCCTTTCAAGAACAGACGCTTTGTTTACAGAAA 793  
QY 122 ---GACGGGAACCTTCTTAAGTCCAGATACAGATGCGAGGAGACACCTCCCTTCCTC 178  
DB 794 GAGGACGGGAACCTTCTTAAGTCCAGATACAGATGCGAGGAGACACCTCCCTTCCTC 853  
QY 179 GTCCTGCTGGTGACCTATCCACAAACAGTTGGCTGAGCGCATGGCCATCCGGCAGACG 238  
DB 854 GTCCTGCTGGTGACCTATCCACAAACAGTTGGCTGAGCGCATGGCCATCCGGCAGACG 913  
QY 239 TGGGGAAAGAGAGGATGTTGAAGGAAAGCAGCTGAAGACATCTTCTCTCTGGGACC 298  
DB 914 TGGGGAAAGAGAGGATGTTGAAGGAAAGCAGCTGAAGACATCTTCTCTCTGGGACT 973  
QY 299 ACCAGCAGTGCAGCGGAACACGAAGAGGTGGACCGAGAGCCGACGACACGGGACATT 358  
DB 974 ACCAGCAGTGCAGCGGAATGAAGAGGTGGACCGAGAGCCGACGACACGGGACATT 1033  
QY 359 ATCCAGAGAGGATTTCTAGACGCTATTACATCTACCTGACCTGAGACCATGATGGGATA 418  
DB 1034 ATCCAGAGAGGATTTCTAGACGCTATTACATCTACCTGACCTGAGACCATGATGGGATG 1093  
QY 419 GAATGGGTCCATCGCTTTGTCTCAGGGCGGCTTTGTGATGAAACAGACATCAGACATG 478  
DB 1094 GAATGGGTCCATCGCTTTGTCTCAGGGCGGCTTTGTGATGAAACAGACATCAGACATG 1153  
QY 479 TTCATCAATGTTGACTATCTGACTGAACCTGCTTCTGAAGAAAAACAGAACACAGGTTT 538  
DB 1154 TTCATCAATGTTGACTATCTGACTGAACCTGCTTCTGAAGAAAAACAGAACACAGGTTT 1213  
QY 539 TTCAGTGCCTTCTGAACCTCAATGAGTTTCCCATCAGGAGCCATTCAGCAAGTGGTTT 598  
DB 1214 TTCAGTGCCTTCTGAACCTCAATGAGTTTCCCATCAGGAGCCATTCAGTAAAGTGGTTT 1273  
QY 599 GTCAGTAATCTGAATATCCGTTGGACAGGTACCCACCATCTCTCCCGCAGCGGTAC 658

Db 1274 GTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATCTTCTGCTGGCACTGCCTAC 1333  
QY 659 GTGTTTTCTGGCAGCTGGGAGTCAAGTGTACAAATGCTCTCCAAAGAGCGTCCCATACATT 718  
Db 1334 GTGTTTTCTGGCAGCTGGGAAAGTCAAGTGTACAAATGCTCTCTGAGAGCGTCCCATACATT 1393  
QY 719 AAAGTGAAGACAGTGTGTTGGGGCTCTGCGCTCGAAAGGCTGAACATCAGATTGGAGGAG 778  
Db 1394 AAAGTGAAGACAGTGTGTTGGGGCTCTGCGCTCGAAAGGCTGAACATCAGATTGGAGGAG 1453  
QY 779 CTCACCTCCAGCGGACGCTTTTTCAGGGGCTTACGCTTCTCGATATGCTCTCTTCAGG 838  
Db 1454 CTCACCTCCAGCGGACGCTTTTTCAGGGGCTTACGCTTCTCGATATGCTCTCTTCAGG 1513  
QY 839 AGGATCGTGGCCTGCCACTTCATCAAGCCTCGGACTCTCTTGGACTACTGGCAGGCTCTA 898  
Db 1514 AGGATCGTGGCCTGCCACTTCATCAAGCCTCGGACTCTCTTGGACTACTGGCAGGCTCTA 1573  
QY 899 GAGAAAT 904  
Db 1574 GAGAAAT 1579  
RESULT 14  
LOCUS AF254738 927 bp DNA linear ROD 24-JUL-2000  
DEFINITION Mus musculus beta 1,3-galactosyltransferase-V (B3gt5) gene, complete cds.  
ACCESSION AF254738  
VERSION AF254738.1 GI:9246994  
KEYWORDS  
SOURCE Mus musculus.  
ORGANISM Mus musculus.  
REFERENCE 1 (bases 1 to 927)  
AUTHORS Zhou,D., Henion,T.R., Jungalwala,F.B., Berger,E.G. and Hennet,T.  
TITLE The beta 1,3-galactosyltransferase beta 3galT-V is a stage-specific embryonic antigen-3 (SSEA-3) synthase  
J. Biol. Chem. 275 (30), 22631-22634 (2000)  
PUBMED 10837462  
REFERENCE 2 (bases 1 to 927)  
AUTHORS Zhou,D. and Hennet,T.  
TITLE Direct Submission  
JOURNAL Submitted (11-APR-2000) Physiology, University of Zurich, Winterthurerstrasse 190, Zurich 8057, Switzerland  
FEATURES  
source Location/Qualifiers  
1..927  
/organism="Mus musculus"  
/strain="129/SvJ"  
/db\_xref="taxon:10090"  
<1..>927  
/gene="B3gt5"  
<1..>927  
/gene="B3gt5"  
/product="beta 1,3-galactosyltransferase-V"  
1..927  
/gene="B3gt5"  
/codon\_start=1  
/product="beta 1,3-galactosyltransferase-V"  
/protein\_id="AAF86241.1"  
/db\_xref="GI:9246995"  
/translation="MAHMKTRLVYASILMMGALCLYFSMDSFRELPPFVKKSHGKFLQ IPDIDCKQPPFLVLLVTSQKQLAARMAIKRTWGRETSVQGGQVRTFFLLGTSSTLE EMATLLESEQRHDIQKDFLDVYNNLTMTMGMEWYHFCQPTAYVMKTDSDMFVN VGYLLELLKNNKTRTFEFTGLKLNLPQPFKSWFVSEYPMWDRYPFPCSGTGYV FSSDVAIQVYNSVYPYIKLEDVFGVGLCLAKLIRPELHTKQTFPFGGLRFSVCRF QXIVACHEMKPDLLTYWQALENSKEDCCPAV"  
BASE COUNT 226 a 241 c 235 g 225 t  
ORIGIN  
Query Match 55.6%; Score 517.8; DB 10; Length 927;

Best Local Similarity 73.8%; Pred. No. 6.3e-130;		Matches 689; Conservative 0; Mismatches 232; Indels 12; Gaps 2;	
Qy	2	ATGGGTTTCCCGAAGATGATGATATATTTGGCTTCTGGTTCTGGGGGCTTTGTT	61
Db	1	ATGGCTACATGAAGACAAGGCTAGTTTACGCTCCATTCATCATGATGGGCGCATTCGC	60
Qy	62	TGTATTTAGCATGTACAGTCTAAATCTTTTCAAAACAGTCTTTGTTTACAAGAA	121
Db	61	TGTATTTAGCATG-----GATCTTTTCAAGAACTCCCGTTTGTTTTAAAGAA	111
Qy	122	G---ACGGGAATCTTAACTCCAGATACAGATGCGAGCGACACACCTCCCTTCCTC	178
Db	112	AGTCAGGGAAGTCTTCAGATTCGGATATAGACTGCAAGCAGAGCCGCTTTCCTG	171
Qy	179	GTCGCTGCTGATGACCTCCCAACAACAGTTGGCTGAGCGCATGCGCATCGGACAGC	238
Db	172	GTCGCTGCTGATGACCTCAACAAGCAGTGGCCGCTGCGATGGCATCCGCAAGACG	231
Qy	239	TGGGGAAAGAGAGGATGTTGAAGGAAACAGCTGAAGACATTTCTCTCTGGGGACC	298
Db	232	TGGGGTAGAGAGACATCTGTGACGGGCCAACAGGTGAGGACCTTCTCTCTGGGGACC	291
Qy	299	ACCAGAGCTCGACGGGAAACGAAGAGGTGGACAGGAGGCCACGACAGCGGGACAT	358
Db	292	TCCGACAGCAGGAGATGGAGCCACAACTCGGAGAGCAGCAGCACCGGACAT	351
Qy	359	ATCCAGAAGATTTCTAGAGCTTATTAACTGACCTGAAAGACCATGATGGGCATA	418
Db	352	ATCCAGAAGACTTCAAGATGCTACTTCAACTGACCTGAAAGACCATGATGGGTATG	411
Qy	419	GAATGGTCCATCGCTTTGTCTCTAGCGCGCTTTCTGATGAACACAGCTCAGACATG	478
Db	412	GAATGGGTACACATTTGTCTCTAGACGCTTACGTGATGAACACGACTCTACATG	471
Qy	479	TTCTCAATGTTGACTATCTGACTGACTGCTTCTGAAGAAAAACAGAACACAGGTTT	538
Db	472	TTTGTGAATGTTGGCTATCTGACGGAACTGCTGCTAAAGAAAAACAAACACAGGTT	531
Qy	539	TTCACTGGCTTCTTGAACCTCAATGATTTCCCATCAGCGACCATTCAGCAAGTGGTT	598
Db	532	TTCAAGGCTACATAAAGCCCGACGACTTTCCCATCCGCGAGAGTTTCAACAAGTGGTT	591
Qy	599	GTCAGTAATCTGAATATCGGTGGGACAGTATCCACCATCTCTCGCGGACCGGCTAC	658
Db	592	GTCAGTAATCTGAATATCCCTGGGACAGGTATCCACCATCTCTCGCGGACTGGTAT	651
Qy	659	GTTGTTTCTGGGACGCTGGGAGTCAAGTCAATCTCTCCACAGCGCTCCCATACAT	718
Db	652	GTTCTTTCCAGGACGCTGGGATCCCAAGTATACAACTCTCCAGAGAGGTTCCGTTTAT	711
Qy	719	AAACTGGAAGACGCTGTTTGTGGGCTCTGCTCGAAAGGCTGAACATCAGATTGGAGGAG	778
Db	712	AAGCTGGAGATGTTGTTGTGGGCTCTGCTGGCCAACTAAAGATCCGCGGAGGAG	771
Qy	779	CTCCATCCCGACCGACCTTTTTCAGGGGGGTTTACGCTTCTCCGTAAGCCTCTTCAGG	838
Db	772	CTGCACACCAAAACAGACCTTCTTCCCTGGCGGTTTACGCTTTCCGTTGCGGCTTCAG	831
Qy	839	AGATGCTGGCCTGCCACTTCATCAACGCTCGGACTCTCTTGGACTATCGGAGGCTCTA	898
Db	832	AAATTTGGGCTGCCATTTTATGAAGCCCGACGACCTGCTCACTTACTTGGCAAGCACTG	891
Qy	899	GAGAAATCCCGGGGAGAGATTGTCGCCCTGTC	931
Db	892	GAGAATCTGAAGAACAGAGGACTGTCTGCTGTC	924

RESULT 15  
AC020851  
LOCUS AC020851 195900 bp DNA linear HTG 15-JUL-2000  
DEFINITION Mus musculus clone RP21-467L12, WORKING DRAFT SEQUENCE, 33  
unordered pieces.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AC020851  
AC020851.2 GI-9211212  
HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
Mus musculus.  
Mus musculus  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 196900)  
DOE Joint Genome Institute.  
Sequencing of Mouse  
Unpublished  
2 (bases 1 to 196900)  
DOE Joint Genome Institute.  
Direct Submission  
Submitted (10-JAN-2000) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
On Jul 15, 2000 this sequence version replaced gi:5686457.  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: http://www.jgi.doe.gov  
-----  
Project Information  
Center Project Name: 1437240  
Center clone name: RPCI-21\_467L12  
-----

Summary Statistics  
Consensus quality: 164618 bases at least Q40  
Consensus quality: 180887 bases at least Q30  
Consensus quality: 183777 bases at least Q20  
Estimated insert size: 147000; pulse field gel estimation  
Estimated insert size: 193700; sum-of-ctnigs estimation  
Quality coverage: 5.67 in Q20 bases; pulse field gel estimation  
Quality coverage: 4.3 in Q20 bases; sum-of-ctnigs estimation.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 33 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1541: contig of 1541 bp in length  
\* 1542 1641: gap of unknown length  
\* 1642 3492: contig of 1851 bp in length  
\* 3493 3592: gap of unknown length  
\* 3593 4738: contig of 1146 bp in length  
\* 4739 4838: gap of unknown length  
\* 4839 7248: contig of 2410 bp in length  
\* 7249 7349: contig of 1830 bp in length  
\* 7349 9179: gap of unknown length  
\* 9179 11493: contig of 2215 bp in length  
\* 11494 11593: gap of unknown length  
\* 11594 13607: contig of 2014 bp in length  
\* 13608 13707: gap of unknown length  
\* 13708 15946: contig of 2239 bp in length  
\* 15947 16046: gap of unknown length  
\* 16047 18392: contig of 2346 bp in length  
\* 18393 18492: gap of unknown length  
\* 18493 21231: contig of 2739 bp in length  
\* 21232 21331: gap of unknown length  
\* 21332 24662: contig of 3331 bp in length  
\* 24663 27140: contig of 2378 bp in length  
\* 27141 29950: contig of 2710 bp in length  
\* 29951 30051: gap of unknown length  
\* 30051 32245: contig of 2195 bp in length  
\* 32246 32345: gap of unknown length  
\* 32346 35688: contig of 3343 bp in length  
\* 35689 40387: gap of unknown length  
\* 40387: contig of 4599 bp in length

* 40388	40487: gap of unknown length	
* 40488	43380: contig of 3093 bp in length	
* 43581	43680: gap of unknown length	
* 43681	46437: contig of 2757 bp in length	
* 46438	46537: gap of unknown length	
* 46538	50184: contig of 3647 bp in length	
* 50185	50284: gap of unknown length	
* 50285	54572: contig of 4288 bp in length	
* 54573	54762: gap of unknown length	
* 54673	60227: contig of 5555 bp in length	
* 60228	60327: gap of unknown length	
* 60328	68024: contig of 7697 bp in length	
* 68025	68124: gap of unknown length	
* 68125	78610: contig of 10486 bp in length	
* 78611	78710: gap of unknown length	
* 78711	88151: contig of 9441 bp in length	
* 88152	88251: gap of unknown length	
* 88252	96257: contig of 8006 bp in length	
* 96258	96357: gap of unknown length	
* 96358	106227: contig of 9870 bp in length	
* 106228	106327: gap of unknown length	
* 106328	115795: contig of 9468 bp in length	
* 115796	115995: gap of unknown length	
* 115896	126363: contig of 10468 bp in length	
* 126364	126463: gap of unknown length	
* 137189	137189: contig of 10726 bp in length	
* 137190	137289: gap of unknown length	
* 146378	146378: contig of 9089 bp in length	
* 146379	146478: gap of unknown length	
* 146479	161686: contig of 15208 bp in length	
* 161687	161786: gap of unknown length	
* 161787	176149: contig of 14363 bp in length	
* 176150	176249: gap of unknown length	
* 176250	196900: contig of 20651 bp in length.	
FEATURES		
source		
1..196900		
/organism="Mus musculus"		
/db_xref="taxon:10090"		
/clone="Rp21-467L12"		
/clone_lib="RPDI mouse PAC library 21"		
BASE COUNT	51456 a 43823 c 44608 g 53666 t 3347 others	
ORIGIN		
Query Match 55.6%; Score 517.8; DB 2; Length 196900;		
Best Local Similarity 73.8%; Pred. No. 7e-130;		
Matches 689; Conservative 0; Mismatches 232; Indels 12; Gaps 2;		
Qy 2	ATGGGTTTCCCGAAGATGAGATTGATGTATATTGCTTCTGTTCTGGGGGCTCTTTGT 61	
Db 102458	ATGGCTACATGAAGACAAGGCTAGTTTACGGCTCCATTCTCATGATGGGCGCACTGTC 102517	
Qy 62	TGTTATTTTAGCATGTACAGTCTAAATCTTTCAAGAACAGTCTTTGTTTACAAGAAA 121	
Db 102518	TGTTACTTCAGCATG-----GATTCTTTCAGAGAAGCTCCGGTTTGTTTTAAAGAAA 102568	
Qy 122	G--*ACGGGAAGTCTCCTTAAGTCCAGATACAGACTGCAGGACAGACCTCCCTTCCTC 178	
Db 102569	AGTCACGGGAAGTTCCTTCAGATTCCGGATATAGACTGCAAGACAGACCGCCCTTTCCTG 102628	
Qy 179	GTCTCTGCTGGTGACCTCATCCCAACAACAGTTGGCTCAGCGCATGGCCATCCGGCAGACG 238	
Db 102629	TGCTGTGCTGGTGACGTCATCTCACAGCAGCTGGCCCTCGCATGGCCATCCGCAAGACG 102688	
Qy 239	TGGGGGAAGAGAGGATGTTGAGGGAAGCAGCTGAAGACATTTCTCTCTGGGGACC 298	
Db 102689	TGGGGTAGAGACATCTGTGCAAGGCCAACAGTGAGGACCTTCTTCTCTGGGGACC 102748	
Qy 299	ACCACAGTGCAGCGGAACGAAGAGGTGGACACAGAGAGCCAGCAGCGGGACATT 358	
Db 102749	TCCGACAGCACCAGGAGATGGACGCCAACACCTCGAGAGCGAGCAGCACCAGCATT 102808	
Qy 359	ATCCAGAGGATTTCTAGACGTCATTACAATCTGACCTGAGACCATGATGGGCATA 418	

Db 102809	ATCCAGAAGGACTTCAAGGATGCCTACTTCAACCTGACCCCTGAAGACCATGATGGGTATG 102868	
Qy 419	GAATGGGTCCATCGCTTTTGTCTTCAGCGGGCTTTGTGATGAAAAACAGACTCAGACATG 478	
Db 102869	GAATGGGTCTACCACCTTTTGTCTTCAGACAGCTTAGTGTGATGAAAAACGGACTCTGACATG 102928	
Qy 479	TTCAATCAATGTTGACTACTCTGACTGAATGCTTCTGNAAGAAAAACAGAACACAGGTTT 538	
Db 102929	TTTGTGAATGTTGGCTATCTGACGGAAGTCTGCTAAAGAAAAACAAAGACGACCATTC 102988	
Qy 539	TTCACTGCTTCTTTGAAACTCAATGAGTTTCCCATCAGCAGCCATTTCAGCAAGTGTTT 598	
Db 102989	TTCAAGGCTACATAAAGCCCCACGACTTCCCATCCGGCAGAAAGTTCAACAAGTGTTT 103048	
Qy 599	GTCAAGTAAATCTGAATATCCGTGGGACAGGTACCCACCACTTCTGCTCCGGCACCGGCTAC 658	
Db 103049	GTGAGTAAGTTCGAATATCCCTGGACAGGTACCCACCTTTTGTCTCCGCTACTGGTTAT 103108	
Qy 659	GTGTTTCTGGCGACGTGGCGAGTCAAGTGTACAATGTCTCCAAGAGCGTCCCATACATT 718	
Db 103109	GTCTTTTCCAGCGACGTGGCGATCCAAGTATACAAGCTCTCAGAGAGCGTTCCGTTTCATC 103168	
Qy 719	AACTGGAAGACGTGTTTGTGGGGCTCTGCCTCGAAAGGCTGAACATCAGATTGGAGGAG 778	
Db 103169	AAGCTGGAGGATGTTTGTGGGCTCTGCTTGGCAAGCTAAAGATCCGCGCGAGGAG 103228	
Qy 779	CTCCACTCCGACGCGACCTTTTCCAGGGGCTTACGCTTCTCGGTATGCTCTTCAGG 838	
Db 103229	CTGCACACCAACACACCTTCTTCCCTGGCGGTTTACGGTTTTTCCGTTGCCGCTTCAG 103288	
Qy 839	AGGATCGTGGCTCGCCACTTTCATCAAGCTCCGACTCTCTTGGACTACTGGCAGGCTCTA 898	
Db 103289	AAAATTTGGCATGCCATTTTATGAAGCCCCAGGACCTGCTCACTTACTGGCAAGCACTG 103348	
Qy 899	GAGAAATCCCGGGGGGAAGATTGTCGGCTGTC 931	
Db 103349	GAGAACTCGAAAGAACAGGACTGTCTCTGCTGTC 103381	

Search completed: April 11, 2003, 20:17:16  
Job time : 2121.98 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 21:30:39 ; Search time 59.5647 Seconds  
(without alignments)  
13710.183 Million cell updates/sec

Title: US-09-914-152-2\_COPY\_401\_1331

Perfect score: 931  
Sequence: 1 aatggcttccgaagatga.....gggaagattgtccgctgtc 931

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 593429 seqs, 438583890 residues

Total number of hits satisfying chosen parameters: 1186858

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	130.4	14.0	1773	10	US-09-739-451-9
2	130.4	14.0	2095	9	US-09-992-598-208
3	130.4	14.0	2095	9	US-09-989-293A-208
4	130.4	14.0	2095	9	US-10-063-547-35
5	130.4	14.0	2095	9	US-09-989-735-208
6	130.4	14.0	2095	9	US-09-990-444-208
7	130.4	14.0	2095	9	US-09-989-730-208
8	130.4	14.0	2095	9	US-09-990-436-208
9	130.4	14.0	2095	9	US-09-991-181-208
10	130.4	14.0	2095	9	US-09-993-687-208
11	130.4	14.0	2095	9	US-09-989-734-208
12	130.4	14.0	2095	9	US-09-997-653-208
13	130.4	14.0	2095	9	US-10-174-590-161
14	130.4	14.0	2095	9	US-10-176-758-161
15	130.4	14.0	2095	9	US-10-063-616-35
16	130.4	14.0	2095	9	US-10-175-737-161
17	130.4	14.0	2095	9	US-09-993-667-208
18	130.4	14.0	2095	9	US-10-063-502-35
19	130.4	14.0	2095	9	US-10-173-706-161

20	130.4	14.0	2095	9	US-10-175-738-161	Sequence 161, App
21	130.4	14.0	2095	9	US-10-175-752-161	Sequence 161, App
22	130.4	14.0	2095	9	US-10-176-482-161	Sequence 161, App
23	130.4	14.0	2095	9	US-10-176-757-161	Sequence 161, App
24	130.4	14.0	2095	9	US-10-176-913-161	Sequence 161, App
25	130.4	14.0	2095	9	US-10-180-552-161	Sequence 161, App
26	130.4	14.0	2095	9	US-10-180-557-161	Sequence 161, App
27	130.4	14.0	2095	9	US-09-990-438-208	Sequence 208, App
28	130.4	14.0	2095	9	US-09-990-562-208	Sequence 208, App
29	130.4	14.0	2095	9	US-09-997-428-208	Sequence 208, App
30	130.4	14.0	2095	9	US-09-997-666-208	Sequence 208, App
31	130.4	14.0	2095	9	US-10-173-700-161	Sequence 161, App
32	130.4	14.0	2095	9	US-10-174-572-161	Sequence 161, App
33	130.4	14.0	2095	9	US-10-174-579-161	Sequence 161, App
34	130.4	14.0	2095	9	US-10-174-582-161	Sequence 161, App
35	130.4	14.0	2095	9	US-10-174-588-161	Sequence 161, App
36	130.4	14.0	2095	9	US-10-175-739-161	Sequence 161, App
37	130.4	14.0	2095	9	US-10-175-740-161	Sequence 161, App
38	130.4	14.0	2095	9	US-10-175-743-161	Sequence 161, App
39	130.4	14.0	2095	9	US-10-176-488-161	Sequence 161, App
40	130.4	14.0	2095	9	US-10-176-492-161	Sequence 161, App
41	130.4	14.0	2095	9	US-10-176-747-161	Sequence 161, App
42	130.4	14.0	2095	9	US-10-176-750-161	Sequence 161, App
43	130.4	14.0	2095	9	US-10-176-985-161	Sequence 161, App
44	130.4	14.0	2095	9	US-10-176-987-161	Sequence 161, App
45	130.4	14.0	2095	9	US-10-176-991-161	Sequence 161, App

ALIGNMENTS

RESULT 1  
US-09-739-451-9  
; Sequence 9, Application US/09739451  
; Patent No. US20010024813A1  
; GENERAL INFORMATION:  
; APPLICANT: Ebner, Reinhard  
; TITLE OF INVENTION: Dendriac and Brainiac-3  
; FILE REFERENCE: PF464  
; CURRENT APPLICATION NUMBER: US/09/739,451  
; CURRENT FILING DATE: 2000-12-19  
; PRIOR APPLICATION NUMBER: 09/213,364  
; PRIOR FILING DATE: 1998-12-17  
; PRIOR APPLICATION NUMBER: 60/077,687  
; PRIOR FILING DATE: 1998-03-12  
; PRIOR APPLICATION NUMBER: 60/108,928  
; PRIOR FILING DATE: 1998-11-17  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 1773  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-739-451-9

Query Match	14.0%	Score 130.4;	DB 10;	Length 1773;
Best Local Similarity	48.9%	Pred. No. 5.6e+33;		
Matches	387;	Conservative	0;	Mismatches 396;
				Indels 9;
				Gaps 1;
QY	111	TTTACAGAAGACGGGAACTTCTTAAGCTCCCGATACAGACTCCAGGCACACCTC	170	
DB	157	TTTACAGAAGACTTTCACTTCACACTTCGAGACATTCAACTCTCATCAAATC	216	
QY	171	CTTCTCGTCTGCTGAGCTCATCCCAACAAAGTTGGTGAGCGCATGGCCATCC	230	
DB	217	CATTCTGGTCACTTCTGGTGACCTCCACCTTCAGATGTGAAGCCAGCGCCATTA	276	
QY	231	GGCAGACGTGGGGAAAGAGAGAGATGGTGAAGGAAAGAGAGATCTTCTCTCC	290	
DB	277	GAGTTACTTGGGTGAAAAAAGTCTTGGTGGGGATATGAGGTTCTTACATTTTCTAT	336	
QY	291	TGGGGACCAACAGCAGTGCAGCGGAAAGAGG-----TGGACCAAGAGGCC	341	

Db 337 TAGGCCAAGAGGCTGAAAGGAGACAAAATGTTGGCATTGTCCTTAGAGGATGACACC 396  
Qy 342 AGCGACACGGGGACATTATCCAGAGGATTTCCCTAGACGCTATTACAACTCGACCCCGA 401  
Db 397 TTCTTTATGGTGACATAATCCGACACGATTTTGTAGACACATATAATACCTGACCTTGA 456  
Qy 402 AGACCATGATGGCCATAGAAATGGTCCATCGCTTTTGTCTCCAGCGCGGCTTGTGATGA 461  
Db 457 AAACCAATTATGGCATTCAGGTGGTAACTAGTGTGGCCCAATGCCCAAGTACGTAATGA 516  
Qy 462 AAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAAGTGGTCTGGAAGAAA 521  
Db 517 AGACACACACTGATGTTTTCATCAATGTTGACTGAAGTGGTCTTAAACCTTAA 576  
Qy 522 ACAGAAACACCGAGTGTTCCTGAGGCTTTGAACTCAATGAGTTTCCCATCATCGGACGC 581  
Db 577 ACCACTCAGAGAAGTGTTCACAGGTTATCCTCTAATTGATAATTATTCCCTATACAGGAT 636  
Qy 582 CATTCAGCAAGTGGTTGTGCTAGTAATCTGAATATCGTGGGACAGGTACCCACCATCT 641  
Db 637 TTTACCAAAAACCCCATATTTCTTACAGGAGTATCCTTCAAGGTGTTCCTCCATAT 696  
Qy 642 GTCGCGGACCGGCTAGCTGTTTCTCGGACGTTGGGAGTCAGGTGTACATGCTCTCA 701  
Db 697 GAGTGGTGTGGTTATATATATGTCACAGATTTGGTCCAGGATCTATGAATGATGG 756  
Qy 702 AGAGCGTCCCATACATATAACTGGAAGACGTTGTTGGGCTCTGCTCGAAAGGCTGA 761  
Db 757 GTCACGTAAACCCATCAAGTTTGAAGATGTTTATGTCGGGATCTGTTGATTTATTA 816  
Qy 762 ACATCAGATGGAGGAGCTCCATCCAGCGGACCTTTTCCAGGGGCTTACGCTTCT 821  
Db 817 AAGTGAACATTCATATTCACAGAGACACAAATCTTTCTTATATAGATCCATTTGG 876  
Qy 822 CGGTATGCTCTTCAGAGGATGTCGGCTGCCACTCATCAAGCTCGGACTCTCTGG 881  
Db 877 ATGCTGTCAACTGAGAGCTGATGTCAGCCCATGCTTCTTCCAGGAGATCA 936  
Qy 882 ACTACTGGCAGG 893  
Db 937 CTTTGGCAGG 948

RESULT 2

US-09-992-598-208  
; Sequence 208, Application US/09992598  
; Patent No. US20020160384A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: P2730PIC20  
; CURRENT APPLICATION NUMBER: US/09/992,598  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088861

PRIOR FILING DATE:	1998-06-11
PRIOR APPLICATION NUMBER:	60/088876
PRIOR FILING DATE:	1998-06-11
PRIOR APPLICATION NUMBER:	60/089105
PRIOR FILING DATE:	1998-06-12
PRIOR APPLICATION NUMBER:	60/089440
PRIOR FILING DATE:	1998-06-16
PRIOR APPLICATION NUMBER:	60/089512
PRIOR FILING DATE:	1998-06-16
PRIOR APPLICATION NUMBER:	60/089514
PRIOR FILING DATE:	1998-06-16
PRIOR APPLICATION NUMBER:	60/089532
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089538
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089598
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089599
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089600
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089653
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089801
PRIOR FILING DATE:	1998-06-18
PRIOR APPLICATION NUMBER:	60/089907
PRIOR FILING DATE:	1998-06-18
PRIOR APPLICATION NUMBER:	60/089908
PRIOR FILING DATE:	1998-06-18
PRIOR APPLICATION NUMBER:	60/089947
PRIOR FILING DATE:	1998-06-19
PRIOR APPLICATION NUMBER:	60/089948
PRIOR FILING DATE:	1998-06-19
PRIOR APPLICATION NUMBER:	60/089952
PRIOR FILING DATE:	1998-06-19
PRIOR APPLICATION NUMBER:	60/090246
PRIOR FILING DATE:	1998-06-22
PRIOR APPLICATION NUMBER:	60/090252
PRIOR FILING DATE:	1998-06-22
PRIOR APPLICATION NUMBER:	60/090254
PRIOR FILING DATE:	1998-06-22
PRIOR APPLICATION NUMBER:	60/090349
PRIOR FILING DATE:	1998-06-23
PRIOR APPLICATION NUMBER:	60/090355
PRIOR FILING DATE:	1998-06-23
PRIOR APPLICATION NUMBER:	60/090429
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090431
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090435
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090444
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090445
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090472
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090535
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090540
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090542
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090557
PRIOR FILING DATE:	1998-06-24
PRIOR APPLICATION NUMBER:	60/090676
PRIOR FILING DATE:	1998-06-25
PRIOR APPLICATION NUMBER:	60/090678
PRIOR FILING DATE:	1998-06-25
PRIOR APPLICATION NUMBER:	60/090590
PRIOR FILING DATE:	1998-06-25
PRIOR APPLICATION NUMBER:	60/090594
PRIOR FILING DATE:	1998-06-25

2	PRIOR APPLICATION NUMBER:	60/090595
3	PRIOR FILING DATE:	1998-06-25
4	PRIOR APPLICATION NUMBER:	60/090696
5	PRIOR FILING DATE:	1998-06-25
6	PRIOR APPLICATION NUMBER:	60/090862
7	PRIOR FILING DATE:	1998-06-26
8	PRIOR APPLICATION NUMBER:	60/090863
9	PRIOR FILING DATE:	1998-06-26
10	PRIOR APPLICATION NUMBER:	60/091360
11	PRIOR FILING DATE:	1998-07-01
12	PRIOR APPLICATION NUMBER:	60/091478
13	PRIOR FILING DATE:	1998-07-02
14	PRIOR APPLICATION NUMBER:	60/091544
15	PRIOR FILING DATE:	1998-07-01
16	PRIOR APPLICATION NUMBER:	60/091519
17	PRIOR FILING DATE:	1998-07-02
18	PRIOR APPLICATION NUMBER:	60/091626
19	PRIOR FILING DATE:	1998-07-02
20	PRIOR APPLICATION NUMBER:	60/091633
21	PRIOR FILING DATE:	1998-07-02
22	PRIOR APPLICATION NUMBER:	60/091978
23	PRIOR FILING DATE:	1998-07-07
24	PRIOR APPLICATION NUMBER:	60/091982
25	PRIOR FILING DATE:	1998-07-07
26	PRIOR APPLICATION NUMBER:	60/092182
27	PRIOR FILING DATE:	1998-07-09

Query Match 14.08: Score 130.4: DB 9: Length 2095:

Best Local similarity 48.98: Pred. No. 6.2e-33:  
 14.00; Score 100.4; DD 3  
 2act1 match

Best Local Similarity 40.38, Fied: NO: 0.26 33,  
Matches 387: Conservative 0: Mismatches 396: Indels 9: Gaps 1:

Qy	111	TTTACAAGAAAGACGGGA	ACTTCTCTTAAGCTCCAGATAGACAGCTGACGGCAGACACCTC	170
Db	494	TTTACAGACAAGACTTTC	ACTTTCACACTTCGAGAGCATTCAAA	CTGCTCATCAAAAATC 553
Qy	171	CCTTCTCGTCTGCTGGT	GTGACCTCATCCCAAAACAGTTGGCTGAGGGCGATGCCCATCC	230
Db	554	CATTTCCTGGTCAATCT	GTGACCTCCACCTTCAGATGTGAAAGCCAGCGAGGCCATT	613
Qy	231	GGCAGACGTGGGGGAA	AGAGAGGATGGTGAAGGAAAGACAGCTGAAGACATCTTCTTCCTCC	290
Db	614	GAGTTACTTGGGGTGA	AAAAAAGCTTGTGGGGGATATGAGGTCTTACATTTTCTTAT	673
Qy	291	TGGGGACCA	CCAGCAGCTGACGGGAAACGAAAGAGG-----TGACCCAGGAGACCC	341
Db	674	TAGGCCAAGAGCTGA	AAAGGACAAAATGTTGGCATTCTCTCTAGAGGATGAACACC	733
Qy	342	AGCGACACGGGGACAT	TATCCAGAAAGATTTCTCTAGACGCTCTATTACAATCTGACCCCTGA	401
Db	734	TTCTTTATGGTGACAT	ATATCCGACAAGATTTTTTAGACACATATAAACCCTGACCTTGA	793
Qy	402	AGACCATGATGGCATAG	AATGGTCCATCGCTTTTCTCTCAGCGGCGTTTGTGTGA	461
Db	794	AAACCATATTGGCATTC	AGGTGGTAACCTGAGTTTTGCCCCAATGCCAAGTACGTAATGA	853
Qy	462	AAACAGACTCAGACAT	GTTCATCAATGTTGACTATCTGACTGAACCTGCTCTGAAGAAA	521
Db	854	AGACAGACACTGATGT	TTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAA	913
Qy	522	ACAGAACCAACCGTTT	TTTCACTGGCTTCTTGAAACTCAATGAGTTTCCCATCAGGCAGC	581
Db	914	ACCCTCAGAGAAGT	TTTTCACAGGTTATCCTCTAATTGATAATTTCTCCTATAGAGAT	973
Qy	582	CATTTCAGCAAGTGGT	TGTGTCAGTAAATCTGAATATCGGTGGGACAGGTACCCACCATCT	641
Db	974	TTTACCAAAAAACCAT	ATATTTCTACAGSAGTATCCTTTCAAGGTGTTCCTCCCATACT	1033
Qy	642	GCTCGGCACCGCTACG	TGTTTCTTGGGACGTGGGAGTFCAGGTGTACAAATGTCTCCA	701
Db	1034	CGAGTGGGTGGGTAT	ATAATGTCAGAGATTTGGTGCCAGGATCTATGAAATGATGG	1093
Qy	702	AGAGCGTCCCATACAT	TAACTGGAAGAGGTGTTTGTGGGGCTCTGCTCGAAAGGCTGA	761

Dd 1094 GTCACGTAACCCATCAAGTTTGAAGATGTTTATGTCGGGATCTGTTGAATTTATTA 1153  
Qy 762 ACATCAGATTGAGGAGTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCT 821  
Dd 1154 AAGTGAACATTCATATCCAGAGACACAATCTTTTCTTTCTATATAGAACATTCATTGG 1213  
Qy 822 CGGTATGCTCTTCAGGAGGATCGTGCGCTGCCACTTTCATCAAGCCCTCGGACTCTTTGG 881  
Dd 1214 ATGCTGTCAACTGAGACGTGATTCAGCCCATGCTTTTCTTCCAAGGAGATCATCA 1273  
Qy 882 ACTACTGGCAGG 893  
Dd 1274 CTTTTGGCAGG 1285

RESULT 3

US-09-989-293A-208  
; Sequence 208, Application US/09989293A  
; Patent No. US2002017164A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: P2730PIC66  
; CURRENT APPLICATION NUMBER: US/09/989,293A  
; CURRENT FILING DATE: 2001-11-20  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088876  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089105  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089440  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089512  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089514  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089532  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089598  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089599  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089600  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089653  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089907  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089908

## RESULT 4

Query Match 14.08; Score 130.4; DB 9; Length 2095:

APPLICANT: Wood,William I.  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3230R1C1  
CURRENT APPLICATION NUMBER: US/10/063,547  
CURRENT FILING DATE: 2002-05-02  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 170  
SEQ ID NO 35  
LENGTH: 2095  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-063-547-35

Query Match 14.0%; Score 130.4; DB 9; Length 2095;  
Best Local Similarity 48.9%; Pred. No. 6.2e-33;  
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;  
QY 111 TTTACAGAAGACGGGAACCTTCCTTAAGCTCCCAAGATACACACTCCAGCGGACACACCTC 170  
Db 494 TTTACAGAAGACCTTCCTTAAGCTCCCAAGATACACACTCCAGCGGACACACCTC 553  
QY 171 CCTTCCTCGCTCGTGGTACCTTCATCCCAAGATGCTGAGCGGATGGCCATCC 230  
Db 554 CATTCCTCGCTCGTGGTACCTTCATCCCAAGATGCTGAGCGGATGGCCATCC 613  
QY 231 GGCAGAGCTGGGGAAGAGAGAGATGTTGAAGGGAAGAGAGATGTTGAGAGATTTCTCTCC 290  
Db 614 GAGTTACTTGGGTGAAAAAAGTCTTGGTGGGATATGAGGTTCTTACATTTTCTTAT 673  
QY 291 TGGGACACACAGCAGCTCCAGCGGAAAGCAAGAGG-----TGGACAGGAGGCC 341  
Db 674 TAGGCAAGAGGCTGAAAGAGCAAAATGTTGGCATTTGCTTAGAGGATGAACACC 733  
QY 342 AGGACACGGGGACATTTACAGAGAGATTTCTAGAGCTCTATTACAATCTGACCCCTGA 401  
Db 734 TTTTATGTTGACATATCCGACAGAGATTTTATGACACATATATATACCTGACCTTGA 793  
QY 402 AGACATGATGGGCATAGAAATGGGTGCTTCTGCTCAGGCGGCTTTGTGATGA 461  
Db 794 AAACATTTATGGCATTCAGGTGGTGAATGATTTGGCCCAATGCCAAGTACGTAATGA 853  
QY 462 AAACAGACTCAGACATGTTCAATATGTTGACTATCTGACTGAACTGCTTCGAAAGAAA 521  
Db 854 AGACAGACACTGATGTTTCAATCAATCAATGCAATTTAGTGAAGTATCTTTTAAACCTTA 913  
QY 522 ACAGAACACCGGTTTTTCACTGGCTTTTGAAGAACTCAATGATTTCCCATCAGGACG 581  
Db 914 ACCACTCAGAGAGTTTTCACTGGCTTTTGAAGAACTCAATGATTTTCCCTATAGAGAT 973  
QY 582 CATTCAGCAAGTGCTTTGTCTAGTAAATCTGAATATCGGTGGACAGGTTACCCACCATTC 641  
Db 974 TTTACCAAAACCCATATTTCTTACAGAGATATCCCTTCAAGGTGTTCCCTCCATAT 1033  
QY 642 GCTCCGACCGGCTAGCTGTTTCTGCGCAGCTGCGGAGTCAAGTGTACAATGTCTCCA 701  
Db 1034 GCAGTGGTGGGTATATATATGATCCAGAGATTTGGTCCCAAGGATCTATGAATGATGG 1093  
QY 702 AGAGCTCCCATACATTAATCAAGTGAAGACGTTGTTGGGGCTCTGCTCGGAAAGCTGA 761  
Db 1094 GTCACGTAAACCCATCAAGTTTGAAGATGTTTATGTCGGGATCTGTTTGAATTTTAA 1153  
QY 762 ACATCAGATTGAGGAGCTCCATCCCGACGACCTTTTCCAGGGGCTTACGGTTCT 821  
Db 1154 AAGTGACATTCATATTTCCAGAGACACAAATCTTTTCTTATATAGAATCCATTTGG 1213  
QY 822 CCGTATGCTCTTCAGGAGGATCGTGGCTGCCACTTCAATCAAGCTCGGACTCTCTTGG 881  
Db 1214 ATGCTGTCACTGAGAGCTGTGATTCGAGGCCATGGCTTTTCTTCCAGGAGATCATCA 1273  
QY 882 ACTACTGGCAGG 893  
Db 1274 CTTTTTGGCAGG 1285

RESULT 5

US-09-989-735-208  
Sequence 208, Application US/09989735  
Publication No. US20020193299A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnovers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730P1C61  
CURRENT APPLICATION NUMBER: US/09/989,735  
CURRENT FILING DATE: 2001-11-19  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029

7	PRIOR APPLICATION NUMBER: 60/090349
7	PRIOR FILING DATE: 1998-06-23
7	PRIOR APPLICATION NUMBER: 60/090355
7	PRIOR FILING DATE: 1998-06-23
7	PRIOR APPLICATION NUMBER: 60/090429
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090431
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090435
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090444
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090445
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090472
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090535
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090540
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090542
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090557
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090676
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090678
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090690
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090594
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090695
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090696
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090862
7	PRIOR FILING DATE: 1998-06-26
7	PRIOR APPLICATION NUMBER: 60/090863
7	PRIOR FILING DATE: 1998-06-26
7	PRIOR APPLICATION NUMBER: 60/091360
7	PRIOR FILING DATE: 1998-07-01
7	PRIOR APPLICATION NUMBER: 60/091478
7	PRIOR FILING DATE: 1998-07-02
7	PRIOR APPLICATION NUMBER: 60/091544
7	PRIOR FILING DATE: 1998-07-01
7	PRIOR APPLICATION NUMBER: 60/091519
7	PRIOR FILING DATE: 1998-07-02
7	PRIOR APPLICATION NUMBER: 60/091626
7	PRIOR FILING DATE: 1998-07-02
7	PRIOR APPLICATION NUMBER: 60/091633
7	PRIOR FILING DATE: 1998-07-02
7	PRIOR APPLICATION NUMBER: 60/091978
7	PRIOR FILING DATE: 1998-07-07
7	PRIOR APPLICATION NUMBER: 60/091982
7	PRIOR FILING DATE: 1998-07-07
7	PRIOR APPLICATION NUMBER: 60/092182
7	PRIOR FILING DATE: 1998-07-09

	Query Match	14.0%	Score 130.4	DB 9	Length 2095
	Best Local Similarity	48.9%	Pred. No. 6.2e-33		
	Matches 387	Conservative 0	Mismatches 396	Indels 9	Gaps 1
QY	111	TTTACAGAAAGACGGGAAC	TTCCTTAAGTCCCGATACAGACTGCAGGCAGACAC	CTC	170
Db	494	TTTACAGACAAGACTTTC	CACTTCACACTTCGAGAGCATTC	CAACTGCTCTCATCAAAATC	553
QY	171	CTTCTCTGTCCTGCTGGT	GACCTCATCCACAAACAGTTGGCTGAGCGCATGCCAT	CC	230
Db	554	CATTTCTGGTCATCTCTGGT	GACCTCCACACCTTCAGATGTGAAAGCCAGGCAGGCCA	TTA	613
QY	231	GGCAGACGTGGGGGAAGAGAGAGATGGT	GTAAGGAAAGACAGCTGAAGACATTTCTTC	CTCC	290

Db 614 GAGTTACTGGGGTGAAGAAAGCTTGGTGGGGATATGAGTCTTACATTTTCTTAT 673  
QY 291 TGGGACCACAGCAGTGCAGCGGAACGAAGAGG-----TGGACAGGAGGCC 341  
Db 674 TAGCCCAAGAGGTGAAGGAACACAAATGTTGGCATTTGCTTAGAGGATGAACACC 733  
QY 342 AGGCACACGGGGGATTCATCAGAAAGATTTCCCTAGACGCTCTATTACAATCTGACCCCTGA 401  
Db 734 TTTCTTTATGTTGACATTAATCCGACAAAGATTTTGTAGACACATATAATACCTGACCTTGA 793  
QY 402 AGACCATGATGGGCATAGAATGGGTCCATCGCTTTTGTCTCAGCGGGGCTTTGTGATGA 461  
Db 794 AAACCATTAATGGCATTCAGGTGGGTAACTGAGTTTGGCCCAATCCCAAGTACGTAATGA 853  
QY 462 AAACAGACTCAGACATGTTCAATGTTGACTATCTGACTGAATGCTTCTGAAAGAAA 521  
Db 854 AGACAGACACTGATGTTTTCATCAATACTGGCAATTTTASTGAAAGTATCTTTTAAACCTAA 913  
QY 522 ACAGAAACACCGAGTTTTCACTGGCTTCTTGAACCTCAATCAGTTTCCATCAGGCAGC 581  
Db 914 ACCACTCAGAGAAGTTTTTCACAGGTTATCTCTTAATGATAATTTATTCCTATAGAGGAT 973  
QY 582 CATTACAGCAAGTGGTTTGTGCTAAATCTGAATATCCGTGGGACAGGTACCCACCAATCT 641  
Db 974 TTTACCAAAAACCCATATTTCTTACCAGGAGTATCTCTTCAAGGTGTTCCCTCCATACT 1033  
QY 642 GCTCCGGCACCGGTACGTTTCTGCGCAGCGTGGCGAGTCAAGTGTACAATGTCCTCA 701  
Db 1034 GCAGTGGTGGGTATATAATGTCACAGATTTGGTCCCAAGGATCTATGAATGATGG 1093  
QY 702 AGACGCTCCCATACATTAATGAACTGAAGACGTTGTTGTGGGCTCTGCTCGAAAGGCTGA 761  
Db 1094 GTCACGTAACACCCATCAAGTTTGAAGATGTTTATGTCGGGATCTGTTGAATTTATTA 1153  
QY 762 ACATCAGATGGAGGAGCTCCACCCAGCGGACCTTTTCCAGGGGCTTACGCTTCT 821  
Db 1154 AGTGAAATCATATTTCCAGACACACAAATCTTTCTTATAGATFCCATTTGG 1213  
QY 822 CGGTATGCTCTTCAGGAGGATCGTGGCTCCCACTTCATCAAGCCTCGGACTCTCTGG 881  
Db 1214 ATGCTGTCACTGAGACGTTGATTCAGGCCCATGGCTTTCTTCCAGGAGATCATCA 1273  
QY 882 ACTACTGGCAGG 893  
Db 1274 CTTTTGGCAGG 1285

RESULT 6  
US-09-990-444-208  
; Sequence 208, Application US/09990444  
; Publication No. US20020193300A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.

; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: P2730PIC19  
; CURRENT APPLICATION NUMBER: US/09/990,444  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826



;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088858  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088861  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088876  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/089105  
;; PRIOR FILING DATE: 1998-06-12  
;; PRIOR APPLICATION NUMBER: 60/089440  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089512  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089514  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089532  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089538  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089598  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089599  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089600  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089653  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089801  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089907  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089908  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089947  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089948  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089952  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090246  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090252  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090254  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090355  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090431  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090472  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090535  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090540  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090542  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090676  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090678  
;; PRIOR FILING DATE: 1998-06-25

;; PRIOR APPLICATION NUMBER: 60/090690  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090694  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

Query Match 14.08; Score 130.4; DB 9; Length 2095;  
Best Local Similarity 48.9%; Pred. No. 6.2e-33;  
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;

QY 111 TTATCAAGAAAGACGGGAACCTTCCTTAAGCTCCAGATACAGACTCGAGCGACACACTC 170  
DB 494 TTACAGACAGACTTTCACCTTCAGACTCGAGACATTCACACTGCTCTCATCAAAATC 553  
QY 171 CCTTCCTCGTCTGCTGGTGACCTCATCCACAAACAGTTGGGTGAGCGCATGGCCATCC 230  
DB 554 CATTCCTGGTCATCTGGTGACCTCCACCTTCAGATGTGAAGCCAGCGACGCATTA 613  
QY 231 GGCAGACGTGGGGAAGAGAGAGATGGTGAAGGAAAGAGCTGAAGACATCTTCTCTCC 290  
DB 614 GAGTTACTTGGGTGAAAAAAGTCTTGGTGGGATATGAGGTCTTACATTTTCTTAT 673  
QY 291 TGGGACCAACAGCAGCTGCACGGGAAACGAAGAGG-----TGGACCAAGAGACC 341  
DB 674 TAGGCCAAGAGGCTGAAAGGAAGACAAAATGTTGGCATTGTCTTAGAGGATGAACACC 733  
QY 342 AGCGACACGGGGACATTCAGAGAGATTCCTAGACGCTCTATTACATCTGACCTGA 401  
DB 734 TTCTTTATGGTGACATAATCCGACAAGATTTTATAGACACATATAATAACCTGACCTGA 793  
QY 402 AGACATGATGGGCATAGAATGGGTCCATCGCTTTGTCTCGAGCGGGCTTTGTGATGA 461  
DB 794 AAACCATATGGCATTCAGGTGGGTAACTGAGTTTGCCCAATGCCAAGTAGCTAATGA 853  
QY 462 AAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACCTGCTCTCAAGAAA 521  
DB 854 AGACAGACACTGATGTTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTAAACCTAA 913  
QY 522 ACAGAACACCAAGGTTTTTCTACTGGCTTCTTGAAGACTCAATGAGTTTCCATCAGGAGC 581  
DB 914 ACCACTCAGAGAAGTTTTTTCACAGGTTATCTCTTAATGATAATTATTCCTATAGAGAT 973  
QY 582 CATTCAGCAAGTGGTTTGTCTAGTAACTGATATCCGTGGGACAGGTACCCACCATCT 641  
DB 974 TTTACCAAAAAACCCATATTTCTTACCAGAGAGTATCTTTCAGGAGTGTTCCTCCCTACT 1033  
QY 642 GCTCCGGCACCGGCTACGTGTTTCTTGCGGAGCTGCGAGTCAAGTGTACATGTCTCCA 701

Db 1034 GCAGTGGGTTGGTTATATATGTCAGAGATTGGTGCAAGGATCTATGAATGATGG 1093  
QY 702 AGAGGGTCCCATACATTAACCTGGAAGACGCTGTTTGTGGGGCTCTGCCTCGAAAGGCTGA 761  
Db 1094 GTCACGTAACCAACCCATCAAGTGTGAAGATGTTTATGTCGGGATCTGTTTGAATTTATTA 1153  
QY 762 ACATCAGATTGGAGGAGCTCCACTCCACCGGACCTTTTTCAGGGGCTTACGCTTCT 821  
Db 1154 AAGTGAACATTCATATTCAGAGACACAAATCTTTTCTATATAGATCCATTGG 1213  
QY 822 CCGTATGCTCTTCAGGAGGATCGGGCTGCCACTTCATCAAGCCTCGGACTCTCTTGG 881  
Db 1214 ATGCTGTCAACTGAGACGCTGTGATTCAGAGCCCATGGCTTTCTTCCAGGAGATCATCA 1273  
QY 882 ACTACTGGCAGG 893  
Db 1274 CTTTTGGCAGG 1285

## RESULT 7

US-09-989-730-208  
; Sequence 208, Application US/09989730  
; Publication No. US20020197674A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Geritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kijavini, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2730P1C69  
; CURRENT APPLICATION NUMBER: US/09/989,730  
; CURRENT FILING DATE: 2001-11-20  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088876  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089105  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089440  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089512  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089514  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089532  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089598  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089599  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089600  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089653  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089801



APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerriksen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kijavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2730P1C14  
CURRENT APPLICATION NUMBER: US/09/990,436  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: 60/043787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444



1 PRIOR FILING DATE: 1997-10-17  
2 PRIOR APPLICATION NUMBER: 60/065186  
3 PRIOR FILING DATE: 1997-11-12  
4 PRIOR APPLICATION NUMBER: 60/065311  
5 PRIOR FILING DATE: 1997-11-13  
6 PRIOR APPLICATION NUMBER: 60/066770  
7 PRIOR FILING DATE: 1997-11-24  
8 PRIOR APPLICATION NUMBER: 60/075945  
9 PRIOR FILING DATE: 1998-02-25  
10 PRIOR APPLICATION NUMBER: 60/078910  
11 PRIOR FILING DATE: 1998-03-20  
12 PRIOR APPLICATION NUMBER: 60/083322  
13 PRIOR FILING DATE: 1998-04-28  
14 PRIOR APPLICATION NUMBER: 60/084600  
15 PRIOR FILING DATE: 1998-05-07  
16 PRIOR APPLICATION NUMBER: 60/087106  
17 PRIOR FILING DATE: 1998-05-28  
18 PRIOR APPLICATION NUMBER: 60/087607  
19 PRIOR FILING DATE: 1998-06-02  
20 PRIOR APPLICATION NUMBER: 60/087609  
21 PRIOR FILING DATE: 1998-06-02  
22 PRIOR APPLICATION NUMBER: 60/087759  
23 PRIOR FILING DATE: 1998-06-02  
24 PRIOR APPLICATION NUMBER: 60/087827  
25 PRIOR FILING DATE: 1998-06-03  
26 PRIOR APPLICATION NUMBER: 60/088021  
27 PRIOR FILING DATE: 1998-06-04  
28 PRIOR APPLICATION NUMBER: 60/088025  
29 PRIOR FILING DATE: 1998-06-04  
30 PRIOR APPLICATION NUMBER: 60/088026  
31 PRIOR FILING DATE: 1998-06-04  
32 PRIOR APPLICATION NUMBER: 60/088028  
33 PRIOR FILING DATE: 1998-06-04  
34 PRIOR APPLICATION NUMBER: 60/088029  
35 PRIOR FILING DATE: 1998-06-04  
36 PRIOR APPLICATION NUMBER: 60/088030  
37 PRIOR FILING DATE: 1998-06-04  
38 PRIOR APPLICATION NUMBER: 60/088033  
39 PRIOR FILING DATE: 1998-06-04  
40 PRIOR APPLICATION NUMBER: 60/088326  
41 PRIOR FILING DATE: 1998-06-04  
42 PRIOR APPLICATION NUMBER: 60/088167  
43 PRIOR FILING DATE: 1998-06-05  
44 PRIOR APPLICATION NUMBER: 60/088202  
45 PRIOR FILING DATE: 1998-06-05  
46 PRIOR APPLICATION NUMBER: 60/088212  
47 PRIOR FILING DATE: 1998-06-05  
48 PRIOR APPLICATION NUMBER: 60/088217  
49 PRIOR FILING DATE: 1998-06-05  
50 PRIOR APPLICATION NUMBER: 60/088655  
51 PRIOR FILING DATE: 1998-06-09  
52 PRIOR APPLICATION NUMBER: 60/088734  
53 PRIOR FILING DATE: 1998-06-10  
54 PRIOR APPLICATION NUMBER: 60/088738  
55 PRIOR FILING DATE: 1998-06-10  
56 PRIOR APPLICATION NUMBER: 60/088742  
57 PRIOR FILING DATE: 1998-06-10  
58 PRIOR APPLICATION NUMBER: 60/088810  
59 PRIOR FILING DATE: 1998-06-10  
60 PRIOR APPLICATION NUMBER: 60/088824  
61 PRIOR FILING DATE: 1998-06-10  
62 PRIOR APPLICATION NUMBER: 60/088826  
63 PRIOR FILING DATE: 1998-06-10  
64 PRIOR APPLICATION NUMBER: 60/088858  
65 PRIOR FILING DATE: 1998-06-11  
66 PRIOR APPLICATION NUMBER: 60/088861  
67 PRIOR FILING DATE: 1998-06-11  
68 PRIOR APPLICATION NUMBER: 60/088876  
69 PRIOR FILING DATE: 1998-06-11  
70 PRIOR APPLICATION NUMBER: 60/089105  
71 PRIOR FILING DATE: 1998-06-12  
72 PRIOR APPLICATION NUMBER: 60/089440  
73 PRIOR FILING DATE: 1998-06-16  
74 PRIOR APPLICATION NUMBER: 60/089512  
75 PRIOR FILING DATE: 1998-06-16  
76 PRIOR APPLICATION NUMBER: 60/089514  
77 PRIOR FILING DATE: 1998-06-16  
78 PRIOR APPLICATION NUMBER: 60/089532  
79 PRIOR FILING DATE: 1998-06-17  
80 PRIOR APPLICATION NUMBER: 60/089538  
81 PRIOR FILING DATE: 1998-06-17  
82 PRIOR APPLICATION NUMBER: 60/089598  
83 PRIOR FILING DATE: 1998-06-17  
84 PRIOR APPLICATION NUMBER: 60/089599  
85 PRIOR FILING DATE: 1998-06-17  
86 PRIOR APPLICATION NUMBER: 60/089600  
87 PRIOR FILING DATE: 1998-06-17  
88 PRIOR APPLICATION NUMBER: 60/089653  
89 PRIOR FILING DATE: 1998-06-17  
90 PRIOR APPLICATION NUMBER: 60/089801  
91 PRIOR FILING DATE: 1998-06-18  
92 PRIOR APPLICATION NUMBER: 60/089907  
93 PRIOR FILING DATE: 1998-06-18  
94 PRIOR APPLICATION NUMBER: 60/089908  
95 PRIOR FILING DATE: 1998-06-18  
96 PRIOR APPLICATION NUMBER: 60/089947  
97 PRIOR FILING DATE: 1998-06-19  
98 PRIOR APPLICATION NUMBER: 60/089948  
99 PRIOR FILING DATE: 1998-06-19  
100 PRIOR APPLICATION NUMBER: 60/089952  
101 PRIOR FILING DATE: 1998-06-19  
102 PRIOR APPLICATION NUMBER: 60/090246  
103 PRIOR FILING DATE: 1998-06-22  
104 PRIOR APPLICATION NUMBER: 60/090252  
105 PRIOR FILING DATE: 1998-06-22  
106 PRIOR APPLICATION NUMBER: 60/090254  
107 PRIOR FILING DATE: 1998-06-22  
108 PRIOR APPLICATION NUMBER: 60/090349  
109 PRIOR FILING DATE: 1998-06-23  
110 PRIOR APPLICATION NUMBER: 60/090355  
111 PRIOR FILING DATE: 1998-06-23  
112 PRIOR APPLICATION NUMBER: 60/090429  
113 PRIOR FILING DATE: 1998-06-24  
114 PRIOR APPLICATION NUMBER: 60/090431  
115 PRIOR FILING DATE: 1998-06-24  
116 PRIOR APPLICATION NUMBER: 60/090435  
117 PRIOR FILING DATE: 1998-06-24  
118 PRIOR APPLICATION NUMBER: 60/090444  
119 PRIOR FILING DATE: 1998-06-24  
120 PRIOR APPLICATION NUMBER: 60/090445  
121 PRIOR FILING DATE: 1998-06-24  
122 PRIOR APPLICATION NUMBER: 60/090472  
123 PRIOR FILING DATE: 1998-06-24  
124 PRIOR APPLICATION NUMBER: 60/090535  
125 PRIOR FILING DATE: 1998-06-24  
126 PRIOR APPLICATION NUMBER: 60/090540  
127 PRIOR FILING DATE: 1998-06-24  
128 PRIOR APPLICATION NUMBER: 60/090542  
129 PRIOR FILING DATE: 1998-06-24  
130 PRIOR APPLICATION NUMBER: 60/090557  
131 PRIOR FILING DATE: 1998-06-24  
132 PRIOR APPLICATION NUMBER: 60/090576  
133 PRIOR FILING DATE: 1998-06-25  
134 PRIOR APPLICATION NUMBER: 60/090678  
135 PRIOR FILING DATE: 1998-06-25  
136 PRIOR APPLICATION NUMBER: 60/090690  
137 PRIOR FILING DATE: 1998-06-25  
138 PRIOR APPLICATION NUMBER: 60/090694  
139 PRIOR FILING DATE: 1998-06-25  
140 PRIOR APPLICATION NUMBER: 60/090695  
141 PRIOR FILING DATE: 1998-06-25  
142 PRIOR APPLICATION NUMBER: 60/090696  
143 PRIOR FILING DATE: 1998-06-25  
144 PRIOR APPLICATION NUMBER: 60/090862  
145 PRIOR FILING DATE: 1998-06-26  
146 PRIOR APPLICATION NUMBER: 60/090863

```
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match          14.0%; Score 130.4; DB 9; Length 2095;
Best Local Similarity 48.9%; Pred. No. 6.2e-33;
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;

QY 111 TTACAAGAAAGACGGGAACCTTCCCTTAAGCTCCAGATACAGACTGCAGGAGACACCTC 170
DB 494 TTTACAGACAAGACTTTCACCTTCGAGAGCATTCAAACTGCTCTCATCAAAATC 553
QY 171 CTTCTCCTGCTGCTGGTGACCTCATCCACAACAGTTGGCTGAGGGCATGGCCATCC 230
DB 554 CATTTCTGCTGCTGCTGGTGACCTCCACCTTCAGATGTAAGAGCCAGGCGCCATTA 613
QY 231 GGCAGACGTGGGGGAAGAGAGATGGTGAAGGAAGAGAGCTCAAGACATCTTCCTCC 290
DB 614 GAGTTACTTGGGGTGAAAAAGAGCTGTTGGTGGGATATGAGTTCTTACATTTTCTAT 673
QY 291 TGGGGACACACAGAGCTGCGAGCGGAACAGAAAGAGG-----TGGACCGAGAGACC 341
DB 674 TAGGCCAAGAGCTGAAAAGGAAGACAAATGTTGGCATTTGCTTAGAGATGACACACC 733
QY 342 AGCGACAGGGGACATTTATCCAGAAGATTTCTAGACGCTATTACAACTGACCCCTGA 401
DB 734 TTCCTTTATGGTGACATAATCCGACAAGATTTTTTAGACATATAATAAACCTGACCTGA 793
QY 402 AGACCATGATGGGCATAGAATGGTCCATCGCTTTTGTCTCAGGCGGCTTTGTGATGA 461
DB 794 AAACCATATGCGATTCAGTGGGTAACTGAGTTTGGCCCAATGCCAAGTACGTAATGA 853
QY 462 AAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACCTGCTCTGAAGAAA 521
DB 854 AGACAGACACTGATGTTTTCATCAATCACTGGCAATTTAGTGAAGTATCTTTTAAACCTAA 913
QY 522 ACAGAACACCAAGGTTTTTTCATCGGCTTTTGAACATCAATGAGTTTCCCATCAGGCAGC 581
DB 914 ACCACTCAGAGAAGTTTTTTCACAGGTTATCCTCTAATTGATAATTATTCCTATAGAGAT 973
QY 582 CATTCAGCAAGTGGTTTGTGCAATAATCTGAATATCGTGGGACAGATACCCACATCTCT 641
DB 974 TTTACCAAAAACCCATATTTCTTACAGAGATATCTTTTCAAGGTTGCTCCTCCATACT 1033
QY 642 GCTCGGCGACCGGTACGTTTCTTGGGACGCTGCGAGTCAAGTGTACAAATCTCTCCA 701
DB 1034 GCAGTGGGTTGGGTTATATAATGTCAGAGATTTGGTGGCAAGATCTATGAAATGATGG 1093
QY 702 AGACGCTCCCATACATTAACCTGAAGAGAGCTGTTTGTGGGGCTCTGCTCGAAAGGCTGA 761
DB 1094 GTCACGTAACCCCACTCAAGTTTGAAGATGTTTATCTCGGATCTGTTTGAATTTATTA 1153
QY 762 ACATCAGATTGGAGAGCTCCACTCCAGCCGACCTTTTTTCCAGGGGGTTAGCCTTCT 821
DB 1154 AAGTGAACATTCATATTCAGAAAGACAAAATCTTTTCTTCTATATAGATCAATTTGG 1213
QY 822 CCGTATGCCTCTTCAGGAGGATCGTGGCGTGCACCTTCATCAAGCCTCGGACTCTCTGG 881
```

```
DB 1214 ATGCTGTCACTGAGACGTGTGATTGCGAGCCCATGCTTTTCTTCCAAGAGATCATCA 1273
QY 882 ACTACTGCGAGG 893
DB 1274 CTTTTTGGCAGG 1285

RESULT 10
US-09-993-687-208
; Sequence 208, Application US/09993687
; Publication No. US20020198149A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C11
; CURRENT APPLICATION NUMBER: US/09/993,687
; CURRENT FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
```

Qy	111	TTTACAAGAAAGACGGGA	CTTCTTAAGCTCCAGATACAGACTCGAGCAGACACCTC	170
Db	494	TTTACAGACAAGAC	TTTCATTTCGAGAGCAATTCAAACTGCTCTCATCAAAATC	553







QY 582 CATTGACGAGTGGTTTGTGTCAGTAATCTGAATATCGTGGGACAGGTACCCACCATCTCT 641  
Db 974 TTTACCAAAAACCATATTTCTTACGAGGATGATCCCTTCAAGGTGTTCCCTCCATCT 1033  
QY 642 GCTCGGCGACGGCTACGTTGTTTCTGGCGACGTGGCGAGTCAGGTGTACAATCTCTCCA 701  
Db 1034 GCAGTGGTGGTTATATAATGTCAGAGATTTGGTGCCAGGATCTATGAATGATGG 1093  
QY 702 AGAGGTCCTCATACATTAACATGGAAGACGTGTTTGTGGGGCTCTGCTCGAAAGGCTGA 761  
Db 1094 GTCACGTAACCCCATCAAGTTGAAGATGTTTATGTCGGGATCTGTTTGAATTTATTA 1153  
QY 762 ACATCAGATTGGAGAGCTCCACTCCACGCCGACCTTTTTCAGGGGGCTTACGCTTCT 821  
Db 1154 AAGTGAACATTCATATCCAGAAGACACAAATCTTTTCTTATATAGAAATCATTGG 1213  
QY 822 CCGTATGCTCTTCAGGAGGATCGTGGCTGCCACTTCATCAAGCTCGGACITCTCTGG 881  
Db 1214 ATGCTGTCACTGAGAGCTGTGATTGACGCCATGCGCTTTTCTTCCAAGGATCATCA 1273  
QY 882 ACTACTGGCAGG 893  
Db 1274 CTTTGGCAGG 1285

RESULT 12

US-09-997-653-208  
; Sequence 208, Application US/09997653  
; Publication No. US20030008297A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; TITLE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: P2730P1C38  
; CURRENT APPLICATION NUMBER: US/09/997,653  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088876  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089105  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089440  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089512  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089514  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089532  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089598  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089599

1	PRIOR FILING DATE:	1998-06-17
2	PRIOR APPLICATION NUMBER:	60/098600
3	PRIOR FILING DATE:	1998-06-17
4	PRIOR APPLICATION NUMBER:	60/098653
5	PRIOR FILING DATE:	1998-06-17
6	PRIOR APPLICATION NUMBER:	60/098801
7	PRIOR FILING DATE:	1998-06-18
8	PRIOR APPLICATION NUMBER:	60/098907
9	PRIOR FILING DATE:	1998-06-18
10	PRIOR APPLICATION NUMBER:	60/098908
11	PRIOR FILING DATE:	1998-06-18
12	PRIOR APPLICATION NUMBER:	60/098947
13	PRIOR FILING DATE:	1998-06-19
14	PRIOR APPLICATION NUMBER:	60/098948
15	PRIOR FILING DATE:	1998-06-19
16	PRIOR APPLICATION NUMBER:	60/098952
17	PRIOR FILING DATE:	1998-06-19
18	PRIOR APPLICATION NUMBER:	60/099024
19	PRIOR FILING DATE:	1998-06-22
20	PRIOR APPLICATION NUMBER:	60/090246
21	PRIOR FILING DATE:	1998-06-22
22	PRIOR APPLICATION NUMBER:	60/090252
23	PRIOR FILING DATE:	1998-06-22
24	PRIOR APPLICATION NUMBER:	60/090254
25	PRIOR FILING DATE:	1998-06-22
26	PRIOR APPLICATION NUMBER:	60/090349
27	PRIOR FILING DATE:	1998-06-23
28	PRIOR APPLICATION NUMBER:	60/090355
29	PRIOR FILING DATE:	1998-06-23
30	PRIOR APPLICATION NUMBER:	60/090429
31	PRIOR FILING DATE:	1998-06-24
32	PRIOR APPLICATION NUMBER:	60/090431
33	PRIOR FILING DATE:	1998-06-24
34	PRIOR APPLICATION NUMBER:	60/090435
35	PRIOR FILING DATE:	1998-06-24
36	PRIOR APPLICATION NUMBER:	60/090444
37	PRIOR FILING DATE:	1998-06-24
38	PRIOR APPLICATION NUMBER:	60/090445
39	PRIOR FILING DATE:	1998-06-24
40	PRIOR APPLICATION NUMBER:	60/090472
41	PRIOR FILING DATE:	1998-06-24
42	PRIOR APPLICATION NUMBER:	60/090535
43	PRIOR FILING DATE:	1998-06-24
44	PRIOR APPLICATION NUMBER:	60/090540
45	PRIOR FILING DATE:	1998-06-24
46	PRIOR APPLICATION NUMBER:	60/090542
47	PRIOR FILING DATE:	1998-06-24
48	PRIOR APPLICATION NUMBER:	60/090557
49	PRIOR FILING DATE:	1998-06-24
50	PRIOR APPLICATION NUMBER:	60/090676
51	PRIOR FILING DATE:	1998-06-25
52	PRIOR APPLICATION NUMBER:	60/090678
53	PRIOR FILING DATE:	1998-06-25
54	PRIOR APPLICATION NUMBER:	60/090690
55	PRIOR FILING DATE:	1998-06-25
56	PRIOR APPLICATION NUMBER:	60/090694
57	PRIOR FILING DATE:	1998-06-25
58	PRIOR APPLICATION NUMBER:	60/090695
59	PRIOR FILING DATE:	1998-06-25
60	PRIOR APPLICATION NUMBER:	60/090696
61	PRIOR FILING DATE:	1998-06-25
62	PRIOR APPLICATION NUMBER:	60/090862
63	PRIOR FILING DATE:	1998-06-26
64	PRIOR APPLICATION NUMBER:	60/090863
65	PRIOR FILING DATE:	1998-06-26
66	PRIOR APPLICATION NUMBER:	60/091360
67	PRIOR FILING DATE:	1998-07-01
68	PRIOR APPLICATION NUMBER:	60/091478
69	PRIOR FILING DATE:	1998-07-02
70	PRIOR APPLICATION NUMBER:	60/091544
71	PRIOR FILING DATE:	1998-07-01
72	PRIOR APPLICATION NUMBER:	60/091519
73	PRIOR FILING DATE:	1998-07-02
74	PRIOR APPLICATION NUMBER:	60/091626
75	PRIOR FILING DATE:	1998-07-02

```
; PRIOR APPLICATION NUMBER: 60/091533
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091538
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09
```

Query Match 14.0%; Score 130.4; DB 9; Length 2095;

Best Local Similarity 48.9%; Pred. No. 6.2e-33;  
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;

QY	111	TTTACAAAGAACGCGGAAGTCTCCTTAAGCTCCAGATACAGACTGCAGGCAGACACTC	170
DB	494	TTTACAGACAGACGATTTCACCTTCACACTTCGAGAGCATTCAACTGCTCTCATCAAAATC	553
QY	171	CTTCTCTCTCTGCTGGTGACCTCATCCACAAACAGTTGGCTGAGCGCATGCCATCC	230
DB	554	CATTTCGTGTCATTCGTGTGACCTCCACCTTCAGATGTAAAGCCAGGAGGCCATT	613
QY	231	GCGACAGCTGGGGGAAGAGAGGATGGTGAAGGGAAGCAGCTGAAGACATCTTCTCTCC	290
DB	614	GAGTTACTTGGGTGAAAAAAGTCTTGGTGGGATATGAGGTTCTTACATTTTCTTAT	673
QY	291	TGGGAGACCACGACGCTGACGGGAACGAAACAGG-----TGGACGAGGAGACC	341
DB	674	TAGGCCAAGAGGCTGAAAAGGAGACAAATGTGGCAITGTCTTASAGGATGAACACC	733
QY	342	ACGACACGGGGACATATCCAGAAGGATTTCTAGACGCTTATTACAACTGCACCTGA	401
DB	734	TTCTTTATGTCACATAATCCGACAGATTTTTCAGACATATATAACTGCACCTTGA	793
QY	402	AGACCATGATGGGCATAGAAATGGTCCATCGCTTTGTCTCAGGCGCGCTTTGTGATGA	461
DB	794	AAACCATATTGCAATTCAGGTGGTAACTGAGTTTGGCCCAATGCCAAGTACGTAATGA	853
QY	462	AAACGAGCTCAGACATGTTTCATCAATGTTGACTATCTGACTGAAGTCTTCTGAAGAAA	521
DB	854	AGACAGACACTGATGTTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAA	913
QY	522	ACAGAACACCCAGGTTTTTCACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGCAGC	581
DB	914	ACCACCTCAGAGAAGTTTTTCAGAGTTATCCTCTAATTGATATATTCTCTATAGAGAT	973
QY	582	CATTACAGCAAGTGGTTGTTCAGTAATCTGAATATCGGTGGGACAGGTACCCACCATCT	641
DB	974	TTTACCAAAAAACCATATTTCTTACCAGGAGTATCCTTTCAGAGTGTTCCTCCCATACT	1033
QY	642	GCTCGGCACCCGGCTACGTTGTTTTCTGGCGACGTGGCGAGTCAAGTGTCATGTCCTCCA	701
DB	1034	CGATGGGTTGGTTATATAATGTCAGAGATTTGGTGCCAGGATCTATGAATGATGG	1093
QY	702	AGAGGTCCTCCATPACATTAACATGGAGACGTTTTGTGGGGCTCTGCCTCGAAGGCTGA	761
DB	1094	GTCAGTAAAACCCCATCAAGTTTGAAGATGTTATGTCGGGATCTGTTTGAATTTATTAA	1153
QY	762	ACATCAGATTTGGAGAGCTCCACTCCACCGGACCTTTTTTTCAGGGGGCTTACGCTTCT	821
DB	1154	AAGTGAACATTCAATTCAGGAAGACACAAATCTTTTCTTCTATATAGAATCCATTTGG	1213
QY	822	CCGATGCTCTTTCAGGAGGATCGTGGCCTGGCCATTCATCAAGCCTCGGACCTCTCTGG	881
DB	1214	ATGCTGTCAACTGAGACGTGTGATTCGAGCCCATGGCTTTCTTCCAAGGAGATCATCA	1273
QY	882	ACTACTGGCAGG	893
DB	1274	CTTTTGGCAGG	1285

RESULT 13  
US-10-174-590-161  
; Sequence 161, Application US/10174590

; Publication No. US20030008352A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430RIC42  
; CURRENT APPLICATION NUMBER: US/10/174,590  
; CURRENT FILING DATE: 2002-06-18  
; Prior application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 161  
; LENGTH: 2095  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-174-590-161

Query Match 14.0%; Score 130.4; DB 9; Length 2095;  
Best Local Similarity 48.9%; Pred. No. 6.2e-33;  
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;  
QY 111 TTACAGAAAGACGGGACCTTCCTTAAGCTCCAGATACAGACTGCAGGAGACACCTC 170  
DB 494 TTTACAGACAGACTTTTCACCTTCACACTTCGAGAGCATTCACAACTCTCATCAAAATC 553  
QY 171 CTTCTCTGCTGCTGTGTGACCTCATCCACAAAGAGTTGGCTGAGCGCATGCCATCC 230  
DB 554 CATTTCTGTCATTTCTGTCACCTCCACCTTCAGATGTGAAGCCAGCGCCATTA 613  
QY 231 GGCAGACGTGGGGAAAGACAGAGGATGGTGAAGGAAAGAGCTGAAGACATTTCTCTCC 290  
DB 614 GAGTTACTTGGGTGAAAAAAGTCTGTGGGGATATGAGTTCTTACATTTTCTAT 673  
QY 291 TGGGACACACGACAGTGCAGCGGAAAGAGG-----TGGACACGAGAGCC 341  
DB 674 TAGCCCAAGAGGCTGAAAGAGACAAAATGTTGGCAATTTGCTTAGAGGATGAACACC 733  
QY 342 AGCGACACGCGGACATTTATCCAGAGGATTTCTTAGAGCTGTTTACAACTTGACCCCTGA 401  
DB 734 TTTCTTATGTTGACATAATCCGACAGATTTTATAGACACATATATACCTGACCTGA 793  
QY 402 AGACCATGATGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGCGAGTCAAGTGTCTCCA 701  
DB 1034 GCAGTGGTGGGTATATATGTCACAGATTTGGTGCACAGGATCTATGAATGATG 1093  
QY 702 AGACGTCCTCATATAAATCGAAGAGCTGTTTGTGGGGCTCTCCCTCGAAAGGCTGA 761  
DB 1094 GTCACGTAACCCATCAAGTTTGAAGATGTTATGTCGGGATCTGTTGAATTTATTA 1153

QY 762 ACATCAGATTGGAGGAGCTCCACTCCAGCGACCTTTTTCAGGGGGCTTAGCTTCT 821  
DB 1154 AAGTGAACATTCATATTCAGAAAGACACAAATCTTTTCTATATAGAAATCATTTGG 1213  
QY 822 CGGTATGCTCTTCAGGAGGATCGTGGCTGCCACTTCATCAAGCCTCGGACTCTCTGG 881  
DB 1214 ATGCTGTCAACTGAGAGCTGTGATTCAGCCCATGGCTTTCTTCAAGAGATCATCA 1273  
QY 882 ACTACTGGCAGG 893  
DB 1274 CTTTGGCAGG 1285  
RESULT 14  
US-10-176-758-161  
; Sequence 161, Application US/10176758  
; Publication No. US20030008353A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430RIC104  
; CURRENT APPLICATION NUMBER: US/10/176,758  
; CURRENT FILING DATE: 2002-06-21  
; Prior application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 161  
; LENGTH: 2095  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-176-758-161

Query Match 14.0%; Score 130.4; DB 9; Length 2095;  
Best Local Similarity 48.9%; Pred. No. 6.2e-33;  
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;  
QY 111 TTACAGAAAGACGGGAACTTCCTTAAGCTCCAGATACAGACTGCAGGAGACACCTC 170  
DB 494 TTTACAGACAGACTTTTCACCTTCACACTTCGAGAGCATTCACAACTCTCATCAAAATC 553  
QY 171 CTTCTCTGCTGCTGTGTGACCTCATCCACAAAGAGTTGGCTGAGCGCATGCCATCC 230  
DB 554 CATTTCTGTCATTTCTGTCACCTCCACCTTCAGATGTGAAGCCAGCGCCATTA 613  
QY 231 GGCAGACGTGGGGAAAGACAGAGGATGGTGAAGGAAAGAGCTGAAGACATTTCTCTCC 290  
DB 614 GAGTTACTTGGGTGAAAAAAGTCTTTGGTGGGATATGAGTTCTTACATTTTCTAT 673  
QY 291 TGGGACACACGACAGTGCAGCGGAAAGAGG-----TGGACACGAGAGCC 341  
DB 674 TAGCCCAAGAGGCTGAAAGAGACAAAATGTTGGCAATTTGCTTAGAGGATGAACACC 733  
QY 342 AGCGACACGCGGACATTTATCCAGAGGATTTCTTAGAGCTGTTTACAACTTGACCCCTGA 401  
DB 734 TTTCTTATGTTGACATAATCCGACAGATTTTATAGACACATATATACCTGACCTGA 793  
QY 402 AGACCATGATGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGCGGCGTTTGTGATGA 461  
DB 794 AAACCATTTATGGCATTCAGGTGGTAACTGAGTTTGGCCCAATGCCAAGTACGTAATGA 853  
QY 462 AAACAGACTCAGACATGTTTCTCAATGTTGACTATCTGACTGAACCTGCTTCTGAAGAAA 521  
DB 854 AGACAGACACTGATGTTTCTCATCAATACTGCGCAATTTAGTGAAGTATCTTTTAAACCTAA 913  
QY 522 ACAGAAACACAGGTTTTCACCTGCTTCTGAACTCAATGAGTTCCCATCAGGAGC 581  
DB 914 ACCACTCAGAGAAGTTTTCACAGGTTATCCCTCAATGATATATTTCTATAGAGGAT 973  
QY 582 CATTCAGCAAGTGTGTTCTCAGTAATCTGAATATCCGTGGGACAGTACCCACCATCT 641  
DB 974 TTTACCAAAAACCCATATTTCTTACCAGGAGTATCTCTTCAAGGTGTTCCCTCCATCT 1033  
QY 642 GTCGCGACCGGCTAGCTGTTTCTGCGAGCTGCGGAGTCAAGTGTACAATGTCTCCA 701  
DB 1034 GCAGTGGTGGGTATATATGTCACAGATTTGGTGCACAGGATCTATGAATGATG 1093  
QY 702 AGACGTCCTCATATAAATCGAAGAGCTGTTTGTGGGGCTCTCCCTCGAAAGGCTGA 761  
DB 1094 GTCACGTAACCCATCAAGTTTGAAGATGTTATGTCGGGATCTGTTGAATTTATTA 1153

Db	854	AGACGAGACACTGATGTTTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAA	913
Qy	522	ACAGAACAAACAGGTTTTTTCACCTGGCTTCTTGAAACTCAATGAGTTTCCCATCAGGAGC	581
Db	914	ACCACTCAGAGAAAGTTTTTCACAGGTATCCCTCAATTGATAAATTTCTCTATAGAGGAT	973
Qy	582	CATTGACAAAGTGTTTTGTCAGTAATCTGAATATCCGTGGGACAGTACCCACCATTCT	641
Db	974	TTTACAAAAAACCCATATTTCTTACCAGGAGTATCCCTTCAAGGTGTTCCCTCCCTACT	1033
Qy	642	GCTCGGCACCGGCTACGTGTTTTCTGGCGAGTGGCGAGTCAGGTGTACAATGTCTCCA	701
Db	1034	GCAGTGGGTTGGGTTATATAATGTCACAGATTTGGTGCCAGGATCTATGAATGATGG	1093
Qy	702	AGAGCGTCCCATACATATAACTGGAAGACGTGTTGTGGGGCTCTGCCTCGAAAGCTGA	761
Db	1094	GTCAGTAAACCCATCAAGTTTGAGATGTTTATGTCGGGATCTGTTTGAATTTATTA	1153
Qy	762	ACATCAGATTTGGAGAGCTCCACTCCCAGCCGACCTTTTTTCCAGGGGCTTACGCTTCT	821
Db	1154	AAGTGAACATTCAATATCCAGAGACACAAATCTTTCTTCTATAGATCCATTTGG	1213
Qy	822	CCGTATGGCTCTTCAGGAGGATCGTGGCTGCCATTCATCAAGGCTCGAGCTCTCTTGG	881
Db	1214	ATGCTGTCAACTGAGACGTGTGATTCAGGCCCATGGCTTTCTTCCAGGAGATCATCA	1273
Qy	882	ACTACTGGCAGG	893
Db	1274	CTTTTTGGCAGG	1285

```

RESULT 15
US-10-063-616-35
; Sequence 35, Application US/10063616
; Publication NO. US20030013855A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063, 616
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See File wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 35
; LENGTH: 2095
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-616-35

```

	Query Match	14.0%;	Score 130.4;	DB 9;	Length 2095;
	Best Local Similarity	48.9%;	Pred. No. 6.2e-33;		
	Matches 387;	Conservative 0;	Mismatches 396;	Indels 9;	Gaps 1;
QY	111	TTTACAGAAAGACGGAACTTCCTTAAGCTCCAGATACAGACTGACGGAGACACCTC	170		
Db	494	TTTACAGACAAGACTTTCACCTTCACACTTCGAGAGCATTTCAACTGCTCTCTCAAAATC	553		
QY	171	CCTTCCTCGTCCTGGTGACCTCATCCACAACAGTTGGCTGAGCGCATGCCATCC	230		
Db	554	CATTTCTGGTCACTCTGGTGACCTCCACCTTCAGATGTGAAGCCAGGCGGCCATT	613		
QY	231	GGCAGAGCTGGGGAAAGAGAGGATGGTGAAGGGAAACGAGCTCAAGACATCTTCCTCC	290		
Db	614	GAGTTACTTGGGGTGAATAAAGACTTGGTGGGGATATGAGGTTCTTCAATTTTCTTAT	673		

Qy	291	TGGGGACCA	CACGAGCTGC	AGCGGAAACG	AAAGAGG-----TGG	ACCAGGAGACC	341	
Db	674	TAGGCCA	AGAGGCTG	AAAGAGGAAC	AAAAATGTTGGCATT	TGCTTTAGAGGATGAACACC	733	
Qy	342	ACGCACAGG	GGACATTTAC	CAGAGGATTTCC	TAGACGTCTATTACA	TCTGACCTGA	401	
Db	734	TTCTTTAT	TGGTGACATAT	CCGACAGATTTTT	TAGACACATATAAT	AACCTGACCTTGA	793	
Qy	402	AGACCATG	ATGGGCATAG	AATGGTGCCAT	CGCTTTTGTCTCAG	GGGGGTTTTGTGATGA	461	
Db	794	AAACCATT	TGSCATTTC	AGTGGTGA	CTGAGTTTTTGCCCA	TGCGCAAGTACGTAAATGA	853	
Qy	462	AAACAGACT	CAGACATGTT	CTCATGTTG	ACTATCTGACTG	NACTGCTTCTGAAGAAA	521	
Db	854	AGACAGACT	GAATGTTTT	TCATCAATACT	GGCAATTTAGTGA	AGTATCTTTTAAACCTAA	913	
Qy	522	ACAGAAACA	CCAGGTTTTT	CTCTGCTCTT	CGAACTCAATGAG	TGTTCCCATCAGCAGC	581	
Db	914	ACCCTCAG	AGNAGTTTT	TCAGAGTTAT	CCCTCAATGATTA	TATCTCTATAGGAT	973	
Qy	582	CATTCAGCA	AGTGTTTTG	TCAGTAATCTG	ATACCGTGGGAC	AGGTACCCACCAATCT	641	
Db	974	TTTACAAAA	AACCCATATTT	CTTACCAGGAGT	ATCCTTTCAAGSGT	TGTTCCCTCCCATACT	1033	
Qy	642	GCTCGGC	ACCGCTACG	TGTTTTCT	GGGAGTGGCGAGT	CAGGTGTCATGTCTCCA	701	
Db	1034	GCAGTGGT	TGGTTATAT	ATATGTC	CCAGATTTGGTG	CCAAAGGTCTATGA	1093	
Qy	702	AGAGCGT	CCCCATAC	ATTAACT	TGGAGACGTGTTT	TGTGGGGCTCTGCC	761	
Db	1094	GTCAGT	AAAAACCC	ATCAAGTTT	GAAATGTTTATG	TCGGGATCTGTTGA	1153	
Qy	762	ACATCAG	ATTGGAGG	AGTCCACT	CCCGACCGACCTTTT	TTCAGGGGGCTTACGCTTCT	821	
Db	1154	AAGTGA	CAATTCAT	ATTCCAGA	AGACAAATCTTTT	CTCTTATATAGA	1213	
Qy	822	CCGTATG	CGCTCTC	AGGAGGAT	CGTGGCTTG	CACTCAAGCT	CGGACTCTCTTGG	881
Db	1214	ATGCTG	CTCACT	GNAGAG	CTGTGAT	TGGACCCATGGCTTTT	TCTTCCAGGAGATCATCA	1273
Qy	882	ACTACTG	GGCAGG	893				
Db	1274	CTTTT	TGGCAGG	1285				

Search completed: April 12, 2003, 03:33:33  
Job time : 74.5647 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 14:22:18 ; Search time 1528.4 Seconds  
(without alignments)  
15994.713 Million cell updates/sec

Title: US-09-914-152-2\_COPY\_492\_1331

Perfect score: 840  
Sequence: 1 ttcaagaacagctcttgt.....99gaagattgtcgcgtgtc 840

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pin.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	840	100.0	933	9	AF145784 Homo sapi
2	840	100.0	2762	9	AB020337 Homo sapi
3	840	100.0	2775	6	E38419 Novel polyp
4	840	100.0	10562	6	E38420 Novel polyp
5	840	100.0	170121	9	AF064860 Homo sapi
6	840	100.0	340000	9	HS21C080 Homo sapi
7	835.8	99.5	933	9	HS46078 Homo sapi
8	808.8	96.3	1576	9	AB041416 Homo sapi
9	797	94.9	1570	9	AB041415 Pan panis
10	789.2	94.0	1565	9	AB041413 Homo sapi
11	785.6	93.5	1566	9	AB041414 Pan trogl
12	783.6	93.3	1360	9	AB041412 Gorilla g
13	764.8	91.0	1579	9	AB041417 Pongo pyg
14	502.2	59.8	927	10	AF254738 Mus muscu
15	502.2	59.8	196900	2	AC020851 Mus muscu
16	497.4	59.2	175861	2	AC120346 Mus muscu
17	320.8	38.2	65400	2	AC109264 Mus muscu
18	185.2	22.0	917	10	AB039137 Mus muscu
19	185.2	22.0	917	10	AB039140 Mus muscu
20	185.2	22.0	917	10	AB039141 Mus muscu
21	185.2	22.0	917	10	AB039142 Mus muscu
22	183.6	21.9	917	10	AB039136 Mus muscu
23	183.6	21.9	917	10	AB039138 Mus muscu
24	183.6	21.9	917	10	AB039143 Mus spici
25	182	21.7	917	10	AB039134 Mus muscu
26	182	21.7	917	10	AB039135 Mus muscu
27	182	21.7	917	10	AB039139 Mus muscu
28	182	21.7	1175	10	AF029790 Mus muscu
29	182	21.7	219368	10	AC098721 Mus muscu
30	166	19.8	1053	9	AB041409 Pan panis
31	165	19.6	1056	9	AB041410 Gorilla g
32	165	19.6	1059	9	AB041411 Pongo pyg
33	164.4	19.6	1060	9	AB041408 Pan trogl
34	163.4	19.5	1037	9	AB041407 Homo sapi
35	163.4	19.5	1739	6	E07739 cDNA encodi
36	163.4	19.5	202001	9	AC016723 Homo sapi
37	161.8	19.3	1909	9	AF117222 Homo sapi
38	142.8	17.0	996	10	AB039156 Mus muscu
39	142.8	17.0	996	10	AB039158 Mus muscu
40	142.8	17.0	996	10	AB039162 Mus muscu
41	142.8	17.0	996	10	AB039163 Mus spici
42	141.2	16.8	996	10	AB039154 Mus muscu
43	141.2	16.8	996	10	AB039155 Mus muscu
44	141.2	16.8	996	10	AB039157 Mus muscu
45	141.2	16.8	996	10	AB039159 Mus muscu

ALIGNMENTS

RESULT 1	AF145784	Homo sapiens betal,3 galactosyltransferase-V (B3GALT5) gene, complete cds.	933 bp	DNA	linear	PRI 14-NOV-1999
AF145784	LOCUS	AF145784	GI:6409192			
DEFINITION	AF145784	AF145784	GI:6409192			
ACCESSION	AF145784	AF145784	GI:6409192			
VERSION	AF145784	AF145784	GI:6409192			
KEYWORDS	AF145784	AF145784	GI:6409192			
SOURCE	AF145784	AF145784	GI:6409192			
ORGANISM	AF145784	AF145784	GI:6409192			
REFERENCE	AF145784	AF145784	GI:6409192			
AUTHORS	AF145784	AF145784	GI:6409192			
TITLE	AF145784	AF145784	GI:6409192			

QY	541	CCACATTCTGCTCGGACCGGCTACGTTTCTGGGACGTGGCGAGTCAAGTGTAC	600
Db	631	CCACATTCTGCTCGGACCGGCTACGTTTCTGGGACGTGGCGAGTCAAGTGTAC	690
QY	601	AATGTCCTCAAGAGCGTCCCATACATTAACTGGAGACGTGTTGTGGGGCTCTGGCTC	660
Db	691	AATGTCCTCAAGAGCGTCCCATACATTAACTGGAGACGTGTTGTGGGGCTCTGGCTC	750
QY	661	GAAGGGCTGAACATCAGATTGGAGAGCTCCACTCCACGGCGACCTTTTTCACAGGGGC	720

751 GAAAGCTGACATCAGATTGAGAGATCCACTCCAGCGACCTTTTTCAGGGGGC 810  
721 TTACGCTTCTCCGATGCCTTTCAGGAGGATCGTGGCCTGCCACTTCATCAAGCCTCG 780  
811 TTACGCTTCTCCGATGCCTTTCAGGAGGATCGTGGCCTGCCACTTCATCAAGCCTCG 870  
781 ACTCTCTGGACTACTGCGACGCTCTAGAGATTCCCGGGGGGAAGATTGTCGCCCTGTC 840  
871 ACTCTCTGGACTACTGCGACGCTCTAGAGATTCCCGGGGGGAAGATTGTCGCCCTGTC 930

RESULT 2	AB020337	2762 bp	linear	PRI 14-APR-2000
LOCUS	AB020337			
DEFINITION	Homo sapiens mRNA for UDP-Gal:GlcNAc betai, 3-galactosyltransferase 5, complete cds.			
ACCESSION	AB020337			
VERSION	AB020337.1	GI:4835502		
KEYWORDS	UDP-gal:glcNAc betai, 3-galactosyltransferase 5, complete cds.			
SOURCE	Homo sapiens			

ORGANISM	TITLE
homo sapiens adeno-ca carcinoma cell line: C630 203 CUNA TO MIURA.	
homo sapiens	Cloning, expression, and characterization of a novel UDP-galactose:beta-N-acetylglucosamine betal,3-galactosyltransferase (betalgal-T5) responsible for
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Primates; Canarrhini; Hominidae; Homo.	
REFERENCE	
AUTHORS	
Isshiki,S., Togayachi,A., Kudo,T., Nishihara,S., Watanabe,M., Kubota,T., Kitajima,M., Shirashi,N., Sasaki,K., Andoh,T. and Narimatsu,H.	

JOURNAL  
MEDLINE  
REFERENCE  
AUTHORS  
TITLE  
DIRECT SUBMISSION  
JOURNAL

synthesis of type 1 chain in colorectal and pancreatic epithelia  
and tumor cells derived therefrom  
J. Biol. Chem. 274 (18), 12499-12507 (1999)  
99230269  
2 (bases 1 to 2762)  
Isshiki,S., Togayachi,A. and Narimatsu,H.  
Direct Submission  
Submitted (20-NOV-1998) Hisashi Narimatsu, Soka University,  
Institute of Life Science; 1-236, Tangecho, Hachioji, Tokyo  
192-8577, Japan (E-mail:sisshiki@pollnet.or.jp,

```

FEATURES
source
    re1:81-426-91-9466; fax:81-426-91-9315)
    Location/Qualifiers
        1..2762
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="21"
            /map="21q22.3"
            /cell_line="Colo 205"
            /cell_type="Adenocarcinoma"
            1..2762
            /gene="betal_3-GALT_5"

```

exon	1..273 /gene="betal,3-Galt 5" /number=1
exon	274..433 /gene="betal,3-Galt 5" /number=3
exon	434..2762 /gene="betal,3-Galt 5" /number=4
CDS	434..11366

/gene="beta1.3-GalT 5"



```
/codon_start=1
/product="UDP-Gal:GlcNAc betal,3-galactosyltransferase 5"
/protein_id="BAA7664.1"
/db_xref="GI:4835503"
/translation="MAFPMRLMYICLLVLGALCLYPSMYSLNPFKEQSFVYKKDGNF
LKLPDTCRQPPFLVLLVTSKHOLAERMAIROTGKERMVKGKOLKTPFLGCTSS
AAETKEVDQESORHGDIIQKDFLDVYVNLTKTMGIEWHRCPOAAYFNKTSDDMF
INVDYLTALLKKNTRTFEFTGLKNEPIRQPFKWFVSKSEYWDYPPFCGSGTG
YVFGDVAQVTVNKSVDYIKLEDVFGLCLERLNIRLEELHSQPTFFFGGLRFSVC
LFRRTVACHFKPRTLDDYQWALENSRGECDPPV"
2762
polyA_site
/gene="betal,3-GalT 5"
/notes="45 a nucleotides"
BASE COUNT      651 a   703 c   679 g   729 t
ORIGIN
Query Match      100.0%; Score 840; DB 9; Length 2762;
Best Local Similarity 100.0%; Pred. No. 1.1e-214;
Matches 840; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTCAAAGAACAGCTCTTTTACAGAAGAGCGGGAACCTCCTTAAGCTCCAGATACA 60
DB 524 TTCAAAGAACAGCTCTTTTACAGAAGAGCGGGAACCTCCTTAAGCTCCAGATACA 583
QY 61 GACTGACGACACACCTCCCTTCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
DB 584 GACTGACGACACACCTCCCTTCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 643
QY 121 GCTGAGCGCATGGCCATCGGCAGACGCTGGGGGAAGAGAGGATGGTGAAGGGAAGCAG 180
DB 644 GCTGAGCGCATGGCCATCGGCAGACGCTGGGGGAAGAGAGGATGGTGAAGGGAAGCAG 703
QY 181 CTGAAGACATTTCTCTCTCTGGGGACACAGCAGCTGCAGCGGAAAGCAAGAGGTGAC 240
DB 704 CTGAAGACATTTCTCTCTGGGGACACAGCAGCTGCAGCGGAAAGCAAGAGGTGAC 763
QY 241 CAGGAGACCGACGACAGCGGGACATTAATCCAGAAGGATTTCTAGACGCTATTACAAT 300
DB 764 CAGGAGACCGACGACAGCGGGACATTAATCCAGAAGGATTTCTAGACGCTATTACAAT 823
QY 301 CTGACCTTGAGACCATGATGGGCATAGATGGGTCCATCGCTTTTGTCTCAGCGGCG 360
DB 824 CTGACCTTGAGACCATGATGGGCATAGATGGGTCCATCGCTTTTGTCTCAGCGGCG 883
QY 361 TTTGTGATGAAGAACAGCTCAGACATGTTCAATGTTGACTATCTGACTGACTGCTT 420
DB 884 TTTGTGATGAAGAACAGCTCAGACATGTTCAATGTTGACTATCTGACTGACTGCTT 943
QY 421 CTGAAGAAAAACAGAACCAAGTGTCTTCACTGGCTTCTTGAAGCTCAATGATTTCCC 480
DB 944 CTGAAGAAAAACAGAACCAAGTGTCTTCACTGGCTTCTTGAAGCTCAATGATTTCCC 1003
QY 481 ATCAGGAGCCATTTCAGCAAGTGGTTGTGCTAATCTGAATATCCGTGGGACAGTAC 540
DB 1004 ATCAGGAGCCATTTCAGCAAGTGGTTGTGCTAATCTGAATATCCGTGGGACAGTAC 1063
QY 541 CCACCATCTCTCTCGGACCGGCTACGTGTTTCTGGGACGTGGGAGTCAAGGTGAC 600
DB 1064 CCACCATCTCTCTCGGACCGGCTACGTGTTTCTGGGACGTGGGAGTCAAGGTGAC 1123
QY 601 AATGTCTCCAGAGCGTCCCATACATTAATTAAGTGAAGAGCTGTTTGTGGGCTCTGCCCTC 660
DB 1124 AATGTCTCCAGAGCGTCCCATACATTAATTAAGTGAAGAGCTGTTTGTGGGCTCTGCCCTC 1183
QY 661 GAAAGGCTGAACATCAGATTGGAGGAGCTCCACTCCAGCCGACCTTTTTCAGGGGGC 720
DB 1184 GAAAGGCTGAACATCAGATTGGAGGAGCTCCACTCCAGCCGACCTTTTTCAGGGGGC 1243
QY 721 TTAGGCTTCCGATGCTCTTCAGGAGGATCGTGGCTGCGACTTCAATCAAGCCTCGG 780
DB 1244 TTAGGCTTCCGATGCTCTTCAGGAGGATCGTGGCTGCGACTTCAATCAAGCCTCGG 1303
QY 781 ACTCTCTTGACTACTGGCAGGCTCTAGAGAAATCCCGGGGGGAAGATTGTCCGCGCTGC 840
```

```
Db 1304 ACTCTTTGGACTACTGCGAGCTCTAGAGAATCCCGGGGGAAGATTCTCCGCTGTC 1363
RESULT 3
E38419 Novel polypeptide. 2775 bp DNA linear PAT 31-JAN-2002
LOCUS E38419
DEFINITION E38419
ACCESSION E38419
VERSION E38419.1 GI:18626993
KEYWORDS JP 2000245464-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2775)
AUTHORS Narimatsu,H., Isshiki,S., Togayauchi,A. and Sasaki,K.
TITLE Novel polypeptide
JOURNAL Patent: JP 2000245464-A 1 12-SEP-2000;
COMMENT KYOWA HAKKO KOGYO CO LTD
OS Homo sapiens (human)
PN JP 2000245464-A/1
PD 12-SEP-2000
PF 25-FEB-1999 JP 1999047571
PR
PI HISASHI NARIMATSU, SOICHIRO ISSHIKI, AKIRA TOGAYAUCHI, PI
KATSUTOSHI SASAKI
PC C12N15/09,A01K67/027,C12N1/21,C12N5/10,C12N9/10,C12P19/00,PC
C12P21/02,
PC C12P21/08,C12Q1/68,G01N33/53//(C12N1/21,C12R1:185),(C12N5/10,
PC C12R1:91),
PC (C12P21/02,C12R1:185),(C12P21/02,C12R1:91),C12N15/00,C12N5/00,
CC (C12N5/00,C12R1:91)
CDS
FT Key Location/Qualifiers
FT CDS Location/Qualifiers
1..2775
/db_xref="taxon:9606"
BASE COUNT      681 a   698 c   669 g   727 t
ORIGIN
Query Match      100.0%; Score 840; DB 6; Length 2775;
Best Local Similarity 100.0%; Pred. No. 1.1e-214;
Matches 840; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTCAAAGAACAGTCCCTTTGTTTACAGAAGAACAGCGGAACCTCCTTAAGCTCCAGATACA 60
DB 492 TTCAAAGAACAGTCCCTTTGTTTACAGAAGAACAGCGGAACCTCCTTAAGCTCCAGATACA 551
QY 61 GACTGACGAGCAGACACCTCCCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
DB 552 GACTGACGAGCAGACACCTCCCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 611
QY 121 GCTGAGCGCATGCCATCCGCGACAGCTGGGGGAAGAGAGGATGGTGAAGGAAAGCAG 180
DB 612 GCTGAGCGCATGCCATCCGCGACAGCTGGGGGAAGAGAGGATGGTGAAGGAAAGCAG 671
QY 181 CTGAAGACATTTCTCTCTCTGGGACCAACAGCAGTGCAGCGGAAACGAAAGAGGTGAC 240
DB 672 CTGAAGACATTTCTCTCTCTGGGACCAACAGCAGTGCAGCGGAAACGAAAGAGGTGAC 731
QY 241 CAGGAGAGCGGACGACGCGGGACATTTATCCAGAAGGATTTCTTAGAGCTCTATTACAAT 300
DB 732 CAGGAGAGCGGACGACGCGGGACATTTATCCAGAAGGATTTCTTAGAGCTCTATTACAAT 791
QY 301 CTGACCTTGAAGACCATGATGGGCATAGATGGTCCATCGCTTTTGTCTCAGCGGGCG 360
DB 792 CTGACCTTGAAGACCATGATGGGCATAGATGGTCCATCGCTTTTGTCTCAGCGGGCG 851
QY 361 TTTGTGATGAAGAACAGCTCAGACATGTTCAATGTTGACTATCTGACTGACTGCTT 420
```



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 170121) Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T., Park,H.-S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E., Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K., Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R., Patterson,D., Reichwald,K., Rump,A., Schillhabel,M.B., Schudy,A., Zimmermann,W., Rosenthal,A., Kudoh,J., Kawasaki,K., Asakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Shimizu,N., Nordsiek,G., Antonarakis,S.E., Minoshima,S., Schaefer,M., Schoen,O., Desario,A., Hornischer,K., Brandt,P., Bloeker,H., Ramser,J., Beck,A., Klages,S., Hennig,J., Riesselmann,L., Dagand,E., Haaf,T., Wehrmeyer,S., Borzym,K., Gardiner,K., Nizetic,D., Francis,F., Leirach,H., Reinhardt,R. and Yaspo,M.Laure. The DNA sequence of human chromosome 21 Nature 405 (6784), 311-319 (2000) 20289799 10830953 2 (bases 1 to 170121) Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B., Weber,J., Schattevoy,R., Yaspo,M.-L. and Rosenthal,A. Direct Submission Submitted (12-MAY-1998) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany 3 (bases 1 to 170121) Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B., Weber,J., Schattevoy,R., Yaspo,M.-L., Rosenthal,A., Yaspo,M.-L. and Rosenthal,A. Direct Submission Submitted (27-FEB-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany 4 (bases 1 to 170121) Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B., Weber,J., Schattevoy,R., Yaspo,M.-L. and Rosenthal,A. Direct Submission Submitted (05-MAR-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany On Feb 27, 2002 this sequence version replaced gi:3171153. Location/Qualifiers 1..170121 /organism="Homo sapiens" /db\_xref="taxon:9606" /chromosome="21" /map="21q22.3" /clone="PAC 70124" 45184 a 36756 c 37509 g 50672 t BASE COUNT ORIGIN Query Match 100.0%; Score 840; DB 9; Length 170121; Best Local Similarity 100.0%; Pred. No. 1.1e-214; Matches 840; Conservative 0; Mismatches 0; Indels 0; Gaps 0; QY 1 TTCAAAGAACAGCTCTCTTTTACAAAGAACGCGGAACTTCCTTAAGCTCCACATACA 60 Db 90138 TTCAAAGAACAGCTCTCTTTTACAAAGAACGCGGAACTTCCTTAAGCTCCACATACA 90137 QY 61 GACTCAGGACGACACCTCCCTTCCTCGCTCTGCTGACCTCATCCCAACAAACAGTTG 120 Db 90198 GACTCAGGACGACACCTCCCTTCCTCGCTCTGCTGACCTCATCCCAACAAACAGTTG 90257 QY 121 GCTGAGCGCATGGCCATCCGGCAGAGCTGGGGAAAGAGAGGATGGTGAAGGGAAGCAG 180 Db 90258 GCTGAGCGCATGGCCATCCGGCAGAGCTGGGGAAAGAGAGGATGGTGAAGGGAAGCAG 90317 QY 181 CTGAAGACATTTCTCTCTGGGACACACAGAGTGCACGCGGAAAGAGAGGTGGAC 240 Db 90318 CTGAAGACATTTCTCTCTGGGACACACAGAGTGCACGCGGAAAGAGAGGTGGAC 90377 QY 241 CAGGAGACCCAGCAGACCGGGACATTATCCAGAAGGATTTCCTAGACGTTCTATTACAAT 300 Db 90378 CAGGAGACCCAGCAGACCGGGACATTATCCAGAAGGATTTCCTAGACGTTCTATTACAAT 90437

QY 301 CTGACCCCTGAAGACCATGATGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGCGCGC 360 Db 90438 CTGACCCCTGAAGACCATGATGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGCGCGC 90497 QY 361 TTTGTGATGAACACAGACTCAGACATGTTTCATCAATTTGACATCTGACTGAACCTGCTT 420 Db 90498 TTTGTGATGAACACAGACTCAGACATGTTTCATCAATTTGACATCTGACTGAACCTGCTT 90557 QY 421 CTGAAGAAAAACAGAACCAACAGGTTTTTTCACCTGGCTTCTTGAACCTCAATGACTTTCCC 480 Db 90558 CTGAAGAAAAACAGAACCAACAGGTTTTTTCACCTGGCTTCTTGAACCTCAATGACTTTCCC 90617 QY 481 ATCAGGAGCCATTTCAGCAAGTGGTTTTCAGTAAATCTGAATATCCGTGGACAGGTAC 540 Db 90618 ATCAGGAGCCATTTCAGCAAGTGGTTTTCAGTAAATCTGAATATCCGTGGACAGGTAC 90677 QY 541 CCACCATTTCTCTCCGGCACCGGCTACGTGTTTTCTGGGACGCTGGCAGTTCAGTGTAC 600 Db 90678 CCACCATTTCTCTCCGGCACCGGCTACGTGTTTTCTGGGACGCTGGCAGTTCAGTGTAC 90737 QY 601 AATGCTCCAAAGAGGTCCTCCATACATTAACCTGGAAGAGCTGTTTGTGGGCTCTGCCTC 660 Db 90738 AATGCTCCAAAGAGGTCCTCCATACATTAACCTGGAAGAGCTGTTTGTGGGCTCTGCCTC 90797 QY 661 GAAGGCTGACATCAGATTGGAGGAGCTCCACTCCAGCCGACCTTTTTTCCAGGGGC 720 Db 90798 GAAGGCTGACATCAGATTGGAGGAGCTCCACTCCAGCCGACCTTTTTTCCAGGGGC 90857 QY 721 TTACGCTTCTCCGTATGCTCTTCAGGAGGATCGTGGCTGCCACTTCATCAAGCCCTCGG 780 Db 90858 TTACGCTTCTCCGTATGCTCTTCAGGAGGATCGTGGCTGCCACTTCATCAAGCCCTCGG 90917 QY 781 ACTCTTTGGACTACTGCGAGGCTCTAGAGAATTCCTGGGGGGAAGATTGTCGCCCTGTC 840 Db 90918 ACTCTTTGGACTACTGCGAGGCTCTAGAGAATTCCTGGGGGGAAGATTGTCGCCCTGTC 90977 RESULT 6 HS21C080 340000 bp DNA linear PRI 24-MAY-2000 LOCUS Homo sapiens chromosome 21 segment HS21C080. ACCESSION AL163280 AP001735 BA000005 VERSION AL163280.2 GI:7717369 KEYWORDS human. SOURCE human. ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 340000) Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T., Park,H.-S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E., Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K., Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R., Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A., Zimmermann,W., Rosenthal,A., Kudoh,J., Shibusaki,K., Asakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuyama,S., Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G., Hornischer,K., Brandt,P., Schaefer,M., Schoen,O., Desario,A., Hennig,J., Riesselmann,L., Dagand,E., Haaf,T., Wehrmeyer,S., Borzym,K., Gardiner,K., Nizetic,D., Francis,F., Leirach,H., Reinhardt,R. and Yaspo,M.L. Direct Submission Submitted (05-MAY-2000) The Chromosome 21 Mapping and Sequencing Consortium: \* RIKEN Genomic Sciences Center, Human Genome Research Group \* Institute of Molecular Biotechnology, Genome Analysis \* Keio University School of Medicine, Dept. of Molecular Biology \* GBF, Dept. of Genome Analysis \* Max-Planck Institute for Molecular Genetics (addresses see below) The Chromosome 21 Mapping and Sequencing Consortium consists of \* RIKEN Genomic Sciences Center, Human Genome Research Group, \* Sagami-hara 228-8555, Japan,

\* e.mail: sakaki@gsr.riken.go.jp  
\* URL: http://hgp.gsc.riken.go.jp/  
and  
\* Institute of Molecular Biotechnology, Genome Analysis, \*  
Beutenbergstrasse 11, D-07745 Jena, Germany,  
\* e.mail: gscj-submit@genome.imb-jena.de  
\* URL: http://genome.imb-jena.de/  
and  
\* Keio University School of Medicine, Dept. of Molecular Biology, \*  
Tokyo 160-8582, Japan,  
\* e.mail: shimizu@dbm-med.keio.ac.jp  
\* URL: http://adenine.dmb.med.keio.ac.jp/  
and  
\* GBF, Dept. of Genome Analysis,  
Mascheroder Weg 1, D-38124 Braunschweig, Germany, \* e.mail:  
info.genome@gbf.de  
\* URL: http://genome.gbf.de/  
and  
\* Max-Planck Institute for Molecular Genetics,  
Innestrasse 73, D-14195 Berlin, Germany,  
\* e.mail: info-chr21@molgen.mpg.de  
\* URL: http://chr21.rz-berlin.mpg.de/  
Location/Qualifiers  
1. 340000  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="21q22.3"  
/chromosome="21"  
/map="21q22.3"  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.3"  
/clone\_lib="RPC11,3-5 PAC library"  
/note="Accession No. AF121897"  
73410. 243533  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.3"  
/clone="P70124"  
/clone\_lib="RPC11,3-5 PAC library"  
/note="Accession No. AF064860"  
224137. >340000  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.3"  
/clone="P206A10, 3' partial"  
/clone\_lib="RPC11,3-5 PAC library"  
/note="Accession No. AF121782"  
286628. >340000  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.3"  
/clone="BAC-291B3, 3' partial"  
/clone\_lib="BAC library"  
/note="Accession No. AF064857"  
1088. 1566  
/note="LIM4"  
/rpt\_family="LINE/L1"  
/rpt\_type=DISPERSED  
2398. 2495  
/note="(TCTA)n"  
/rpt\_family="Simple\_repeat"  
/rpt\_type=TANDEM  
complement(2496. 2878)  
/note="THE1C"  
/rpt\_family="LTR/MaLR"  
/rpt\_type=DISPERSED  
2879. 2931

FEATURES  
source  
/note="(TCTA)n"  
/rpt\_family="Simple\_repeat"  
/rpt\_type=TANDEM  
2932. 3004  
/note="(TCCA)n"  
/rpt\_family="Simple\_repeat"  
/rpt\_type=TANDEM  
3005. 3074  
/note="(TCCA)n"  
/rpt\_family="Simple\_repeat"  
/rpt\_type=TANDEM  
3473. 19157  
/gene="SH3BGR"  
join(<3473. 3565,12649. 12678,15331. 15460,18739. 19157)  
/gene="SH3BGR"  
join(3473. 3565,12649. 12678,15331. 15426)  
/partial  
/gene="SH3BGR"  
/note="Accession No. X93498"  
/codon\_start=1  
/product="21-Glutamic Acid Rich protein 21-CARP"  
/protein\_id="CAB90445.1"  
/db\_xref="GI:7717370"  
/translation="GSEKAEGETEAKGSESDVGNLPEAQEKNEBEGETATEETEE  
TAMEGAEAEAEETAEGETEPEDEDS"  
3473. 3565  
/gene="SH3BGR"  
/number=4  
5542. 5829  
/note="AluSx"  
/rpt\_family="SINE/Alu"  
/rpt\_type=DISPERSED  
complement(6887. 7067)  
/note="L1MD3"  
/rpt\_family="LINE/L1"  
/rpt\_type=DISPERSED  
complement(8482. 8762)  
/note="AluJo"  
/rpt\_family="SINE/Alu"  
/rpt\_type=DISPERSED  
complement(8865. 8964)  
/note="L2"  
/rpt\_family="LINE/L2"  
/rpt\_type=DISPERSED  
9131. 9424  
/note="AluSq"  
/rpt\_family="SINE/Alu"  
/rpt\_type=DISPERSED  
complement(9723. 9882)  
/note="L2"  
/rpt\_family="LINE/L2"  
/rpt\_type=DISPERSED  
complement(9989. 10077)  
/note="L2"  
/rpt\_family="LINE/L2"  
/rpt\_type=DISPERSED  
complement(10078. 10429)  
/note="THE1B"  
/rpt\_family="LTR/MaLR"  
/rpt\_type=DISPERSED  
complement(10430. 10651)  
/note="L2"  
/rpt\_family="LINE/L2"  
/rpt\_type=DISPERSED  
10980. 11000  
/note="(TTTA)n"  
/rpt\_family="Simple\_repeat"  
/rpt\_type=TANDEM  
complement(11002. 11282)  
/note="AluSp"  
/rpt\_family="SINE/Alu"  
/rpt\_type=DISPERSED  
complement(11315. 11556)

		/note="AluJo"			
		/rpt_family="SINE/Alu"			
		/rpt_type=DISPERSED			
		11702..11753			
		/note="MIR"			
		/rpt_family="SINE/MIR"			
		/rpt_type=DISPERSED			
		12649..12678			
		/gene="SH3BGR"			
		/number=5			
	repeat_region	complement(13115..13425)			
		/note="Alusx"			
		/rpt_family="SINE/Alu"			
		/rpt_type=DISPERSED			
	repeat_region	complement(13598..13892)			
		/note="Aluy"			
		/rpt_family="SINE/Alu"			
		/rpt_type=DISPERSED			
	repeat_region	14013..14131			
		/note="(CA)n"			
		/rpt_family="Simple_repeat"			
		Query Match 100.0%; Score 840; DB 9; Length 340000;			
		Best Local Similarity 100.0%; Pred. No. 1.le-214;			
		Matches 840; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	TTCAAAGAACAGTCCTTTGGTTTACAAGAAGACGCGGAACATTCCTTAAAGTCCCAGATACA	60		
Db	163546	TTCAAAGAACAGTCCTTTGGTTTACAAGAAGACGCGGAACATTCCTTAAAGTCCCAGATACA	163605		
QY	61	GACTGCAGGCAGACACCTCCCTTCCTCGTCTCTGTGTTGACCTCATCCACAAAACAGTTG	120		
Db	163606	GACTGCAGGCAGACACCTCCCTTCCTCGTCTCTGTGTTGACCTCATCCACAAAACAGTTG	163665		
QY	121	GCTGAGCGCATGCCATCCGCGACAGCTGGGGAAAGAGAGGATGGTGAAGGGAAAGCAG	180		
Db	163666	GCTGAGCGCATGCCATCCGCGACAGCTGGGGAAAGAGAGGATGGTGAAGGGAAAGCAG	163725		
QY	181	CTGAAGACATTCTTCCTCTGGGACACCACGACGTGCAGCGGAAACGAAAGAGGTGGAC	240		
Db	163726	CTGAAGACATTCTTCCTCTGGGACACCACGACGTGCAGCGGAAACGAAAGAGGTGGAC	163785		
QY	241	CAGGAGACCCAGGACACGGGACATATCCAGAAGGATTCCTAGACGTCTATTACAAT	300		
Db	163786	CAGGAGACCCAGGACACGGGACATATCCAGAAGGATTCCTAGACGTCTATTACAAT	163845		
QY	301	CTGACCTCGAAGACCATGATGGCATAGAATGGGTCCATCGCTTTTGTCCTCAGGGCGG	360		
Db	163846	CTGACCTCGAAGACCATGATGGCATAGAATGGGTCCATCGCTTTTGTCCTCAGGGCGG	163905		
QY	361	TTTGATGATAAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACGTGCTT	420		
Db	163906	TTTGATGATAAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACGTGCTT	163965		
QY	421	CTGAAGAAAAACAGAACACCCAGTTTTTCACCTGGCTCTTTGAAACTCAATGAGTTTCCC	480		
Db	163966	CTGAAGAAAAACAGAACACCCAGTTTTTCACCTGGCTCTTTGAAACTCAATGAGTTTCCC	164025		
QY	481	ATCAGGCAGCCATTACGCAAGTGGTTTGTCAGTAATACTGAATATCCGTGGACAGGTAC	540		
Db	164026	ATCAGGCAGCCATTACGCAAGTGGTTTGTCAGTAATACTGAATATCCGTGGACAGGTAC	164085		
QY	541	CCACATTCTGCTCGGGCACCGGCTAGTGTCTTCTGCGCAGCTGGCCAGTCAGGTGTAC	600		
Db	164086	CCACATTCTGCTCGGGCACCGGCTAGTGTCTTCTGCGCAGCTGGCCAGTCAGGTGTAC	164145		
QY	601	AATGTCTCCAAGAGGTCCTCATACATTAAACTGGAAGACGTGTTGTGGGGCTCTGCCCTC	660		
Db	164146	AATGTCTCCAAGAGGTCCTCATACATTAAACTGGAAGACGTGTTGTGGGGCTCTGCCCTC	164205		
QY	661	GAAGGCTGAACATCAGATTGGAGGAGCTCCACTCCCAGCGCACCTTTTTCACGGGGGC	720		
Db	164206	GAAGGCTGAACATCAGATTGGAGGAGCTCCACTCCCAGCGCACCTTTTTCACGGGGGC	164265		

QY	721	TTACGGTTCTCCGGTATCGCTCTTCAGGAGGATCGTGGCTTGCCTTCCACTTTCATCAAGCCTCGG	780
Db	164266	TTACGGTTCTCCGGTATCGCTCTTCAGGAGGATCGTGGCTGCCACTTTCATCAAGCCTCGG	164325
QY	781	ACTCTCTTGACACTACTGGCAGGCTCTAGAGAATTCCTCCGGGGGAGAGATTGTTCGGCCTGTC	840
Db	164326	ACTCTCTTGACACTACTGGCAGGCTCTAGAGAATTCCTCCGGGGGAGAGATTGTTCGGCCTGTC	164385
RESULT 7			
HSA6078			
LOCUS			
DEFINITION Homo sapiens beta3gal-T6 gene.			
ACCESSION AJ006078			
VERSION AJ006078.1 GI:7799922			
KEYWORDS beta-1,3-galactosyltransferase; beta3gal-T6 gene.			
SOURCE Homo sapiens.			
ORGANISM Homo sapiens.			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE 1 (bases 1 to 933)			
AUTHORS Amado, M., Carneiro, F. and Clausen, H.			
TITLE Cloning and expression of two beta-1,3-galactosyltransferases: beta3gal-T5 and beta3gal-T6			
JOURNAL Unpublished			
REFERENCE 2 (bases 1 to 933)			
AUTHORS Amado, M.			
TITLE Direct Submission			
JOURNAL Submitted (11-MAY-1998) Amado M., Department of Oral Diagnostics, Royal dental School, Norre Alle 20, 2200 Copenhagen, DENMARK			
FEATURES			
source			
1..933			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
gene			
1..933			
/gene="beta3gal-T6"			
CDS			
1..933			
/gene="beta3gal-T6"			
/codon_start=1			
/product="beta-1,3-galactosyltransferase"			
/protein_id="CA891547.1"			
/db_xref="GI:7799923"			
/translation="MAPKMRIMVICLLVLGALCYYSMSNLNPFKEOSFYVYKKDGNF LKLPDTCRTQTFPLVLLVTSKQLAERMAIRQTGKERTVKGQKLKTLFFLLGTTSS AKETEVDOESQRHGDIIQKDFLDVYVNLTKTMGIEWHRFCPOAAAFVMTKDSDMF INVYDITELLKKNKTRFTFTGFKLNEFFIRQPFKWFVSKSVPMWDRYPFPCSGTG YVFGDVAOVNVNKSVPYIKLEDVFVGLCLERLNIRLELHSQLTPFPGLRPSVC LFRRIVACHFIKPRLLTDVQALENSRGDCPPV"			
BASE COUNT 229 a 234 c 240 g 229 t 1 others			
ORIGIN			
Query Match 99.5%; Score 835.8; DB 9; Length 933;			
Best Local Similarity 99.6%; Pred. No. 1.4e-213;			
Matches 837; Conservative 0; Mismatches 3; Indels 0; Gaps 0;			
Qy	1	TTCAAGAACAGTCCTTTGTTTACAAGAACAGCGGGAACCTTCCTTAAGCTCCCATACA 60	
Db	91	TTCAAGAACAGTCCTTTGTTTACAAGAACAGCGGGAACCTTCCTTAAGCTCCCATACA 150	
Qy	61	GACTGCAGCAGACACCTCCCTTCCTGCTGCTGCTGAGACCTCATCCCAAAACAGTTG 120	
Db	151	GACTGCAGCAGACACCTCCCTTCCTGCTGCTGCTGAGACCTCATCCCAAAACAGTTG 210	
Qy	121	GCTGAGCGCATGCCATCCGGCAGAGCTGGGGGAAAGAGAGGATGGTGAAGGGAAGCAG 180	
Db	211	GCTGAGCGCATGCCATCCGGCAGAGCTGGGGGAAAGAGAGCGGTGAAGGGAAGCAG 270	
Qy	181	CTGAAGACATTCTTCTCCTCTGGGACACACAGCAGTGCAGCGGAAACGAAAGAGGTGGAC 240	
Db	271	CTGAAGACATTCTTCTCCTCTGGGACACACAGCAGTGCAGCGGAAACGAAAGAGGTGGAC 330	
Qy	241	CAGGAGACCCAGCGACACGGGGACATTATCCAGGAAGGATTTCCTAGACGCTATTACAAAT 300	

```
Db 331 CAGGAGAGCCGACGACGGGACATATCCAGAAGGATTTCCAGAGCTTATTACAAAT 390
|||||
Qy 301 CTGACCTTGAAGACCATGATGGCAGATAGATGGGTCCATCGCTTTTGTCTCCAGCGCGG 360
|||||
Db 391 CTGACCTTGAAGACCATGATGGCAGATAGATGGGTCCATCGCTTTTGTCTCCAGCGCGG 450
|||||
Qy 361 TTTGTGATGAAACAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
|||||
Db 451 TTTGTGATGAAACAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 510
|||||
Qy 421 CTGAGAGAAACAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
|||||
Db 511 CTGAGAGAAACAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 570
|||||
Qy 481 ATCAGGACCCATTCAGCAAGTGGTGTGTCAGTAAATCTGAATATCCGTGGGACAGGTAC 540
|||||
Db 571 ATCAGGACCCATTCAGCAAGTGGTGTGTCAGTAAATCTGAATATCCGTGGGACAGGTAC 630
|||||
Qy 541 CCACCATTCGTCCGGCACCAGCTACGTGTTTCTGGCAGCGTGGGAGTCAGGTGTAC 600
|||||
Db 631 CCACCATTCGTCCGGCACCAGCTACGTGTTTCTGGCAGCGTGGGAGTCAGGTGTAC 690
|||||
Qy 601 AATGTCTCAAGAGCGTCCATACATTAATCTGAAGACGTGTTTGTGGGCTCTGCCTC 660
|||||
Db 691 AATGTCTCAAGAGCGTCCATACATTAATCTGAAGACGTGTTTGTGGGCTCTGCCTC 750
|||||
Qy 661 GAAAGCGTGAACATCAGATGGAGGAGCTCCACTCCAGCGGACCTTTTCCAGGGGGC 720
|||||
Db 751 GAAAGCGTGAACATCAGATGGAGGAGCTCCACTCCAGCGGACCTTTTCCAGGGGGC 810
|||||
Qy 721 TTACGCTTCTCGTATGCTCTTCAGGAGGATCGTGCCCTGCCACTTCATCAAGCGCTCGG 780
|||||
Db 811 TTACGCTTCTCGTATGCTCTTCAGGAGGATCGTGCCCTGCCACTTCATCAAGCGCTCGG 870
|||||
Qy 781 ACTCTTTGGACTACTGGCAGGCTCTAGAGATTCGCGGGGGAAGATTTCCGCGCTGTC 840
|||||
Db 871 ACTCTTTGGACTACTGGCAGGCTCTAGAGATTCGCGGGGGAAGATTTCCGCGCTGTC 930
|||||

RESULT 8
LOCUS AB041416 1576 bp DNA linear PRI 17-OCT-2000
DEFINITION Homo sapiens beta1,3-GalT 5 gene for UDP-cal:GlcNAc
beta1,3-galactosyltransferase 5, partial cds.
ACCESSION AB041416
VERSION AB041416.1 GI:7593026
KEYWORDS
SOURCE Homo sapiens (isolate:#056) DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1576)
AUTHORS Liu,Y. and Saitou,N.
JOURNAL Silver Project
TITLE Published Only in DataBase (2000)
REFERENCE 2 (bases 1 to 1576)
AUTHORS Liu,Y. and Saitou,N.
JOURNAL Direct Submission
TITLE Submitted (11-APR-2000) Naruya Saitou, National Institute of
Genetics, Laboratory of Evolutionary Genetics; Ilii Yata, Mishima,
Shizuoka 411-8540, Japan (E-mail:nsaitou@genetics.nig.ac.jp,
URL:http://sayer.lab.nig.ac.jp/-silver/, Tel:81-559-81-6790,
Fax:81-559-81-6789)
FEATURES
Location/Qualifiers
source 1..1576
organism="Homo sapiens"
isolate="#056"
/db_xref="taxon:9606"
/note="human sequence used for primer design based on Acc#
AB020337"
<1..27
/number=3
exon
```

```
intron 28..674
/number=3
gene 675..1576
/gene="beta1,3-GalT 5"
CDS 675..>1576
/gene="beta1,3-GalT 5"
/codon_start=1
/product="UDP-Gal:GlcNAc beta1,3-galactosyltransferase 5"
/protein_id="BAA94501.1"
/db_xref="GI:7593027"
/translation="MAPFKRLMYICLLVLGALCLYFSMYSLNPPKQSFVYKKDNF
LKLPDTCRQTPPELVLLVTSSHQLAERMAIRQWTKERTVKGKLTFFLLGTSS
AATKEVDQSSQRHDIQKDFLDVYNYLTKTMGIEWHVFQPAQAFVPPCSDMF
INVDYLTLLKKNRTFRFFFLKLEFPFQPFQKSWFVSEYQWDRVPPCSDGTG
YVSGDVASOVYNVSKSVPIKLEDFVFLCLERLNRLEELHSQPTFFPGGLRFSVC
LFRIVACHFIKPTLLDYQWALE"
exon 675..>1576
/gene="beta1,3-GalT 5"
/number=4
BASE COUNT 415 a 373 c 387 g 401 t
ORIGIN
Query Match 96.3%; Score 808.8; DB 9; Length 1576;
Best Local Similarity 99.8%; Pred. No. 2.6e-206;
Matches 810; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 TTCAAGAACAGCTCTTCTTTTACAGAAAGACGGGAACCTTCCTTAAGCTCCAGATACA 60
Db 765 TTCAAGAACAGCTCTTCTTTTACAGAAAGACGGGAACCTTCCTTAAGCTCCAGATACA 824
Qy 61 GACTGCGAGGACAGACCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db 825 GACTGCGAGGACAGACCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 884
Qy 121 GCTGAGCGCATGGCCATCCGCGACAGCTGGGGGAAAGAGAGAGGATGCTGAAGGGAAGACAG 180
Db 885 GCTGAGCGCATGGCCATCCGCGACAGCTGGGGGAAAGAGAGAGGATGCTGAAGGGAAGACAG 944
Qy 181 CTGAAGACATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
Db 945 CTGAAGACATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1004
Qy 241 CAGGAGAGCCAGCAGCAGCGGGGACATTTATCCAGAGGATTTCTCTAGAGCTCTATTACAAAT 300
Db 1005 CAGGAGAGCCAGCAGCAGCGGGGACATTTATCCAGAGGATTTCTCTAGAGCTCTATTACAAAT 1064
Qy 301 CTGACCTTGAAGACCATGATGGGATAGAAATGGGTCCATCGCTTTTGTCTCTAGCGCGCG 360
Db 1065 CTGACCTTGAAGACCATGATGGGATAGAAATGGGTCCATCGCTTTTGTCTCTAGCGCGCG 1124
Qy 361 TTTGTGATGAAACAGACATCAGACATGTTTCATCAATTTTGACTATCTGACTGAACTGCTT 420
Db 1125 TTTGTGATGAAACAGACATCAGACATGTTTCATCAATTTTGACTATCTGACTGAACTGCTT 1184
Qy 421 CTGAAGAAACAGACAAACAGGTTTTCCTGCTCTCTTGAAGCTCAATGAGTTTCCC 480
Db 1185 CTGAAGAAACAGACAAACAGGTTTTCCTGCTCTCTTGAAGCTCAATGAGTTTCCC 1244
Qy 481 ATCAGGACCGCATTCAGCAAGTGGTGTGTCAGTAAATCTGAATATCCGTGGGACAGGTAC 540
Db 1245 ATCAGGACCGCATTCAGCAAGTGGTGTGTCAGTAAATCTGAATATCCGTGGGACAGGTAC 1304
Qy 541 CCACCATTCGTCCGGCACCAGCTACGTGTTTCTGGCAGCGTGGGAGTCAGGTGTAC 600
Db 1305 CCACCATTCGTCCGGCACCAGCTACGTGTTTCTGGCAGCGTGGGAGTCAGGTGTAC 1364
Qy 601 AATGTCTCAAGAGCGTCCCATACATTAATCTGAAGACGTGTTTGTGGGCTCTGCCCTC 660
Db 1365 AATGTCTCAAGAGCGTCCCATACATTAATCTGAAGACGTGTTTGTGGGCTCTGCCCTC 1424
Qy 661 GAAAGCGTGAACATCAGATGGAGGAGCTCCACTCCAGCGGACCTTTTCCAGGGGGC 720
Db 1425 GAAAGCGTGAACATCAGATGGAGGAGCTCCACTCCAGCGGACCTTTTCCAGGGGGC 1484
```

Qy 721 TTACGCTTCCTCGTATGCTCTTTAGAGAGGATCGTGCCCTGCCCTCATCAAGCCTCGG 780  
|||||  
Db 1485 TTACGCTTCCTCGTATGCTCTTTAGAGAGGATCGTGCCCTGCCCTCATCAAGCCTCGG 1544  
|||||  
Qy 781 ACTCTCTTGACTACTGGCAGGCTCTAGAGAA 812  
|||||  
Db 1545 ACTCTCTTGACTACTGGCAGGCTCTAGAGAA 1576  
|||||  
RESULT 9  
AB041415 1570 bp DNA linear PRI 13-APR-2000  
LOCUS Pan paniscus betal,3-Galt 5 gene for UDP-Gal:GlcNAC  
DEFINITION betal,3-galactosyltransferase 5, partial cds.  
ACCESSION AB041415  
VERSION AB041415.1 GI:7593024  
KEYWORDS  
SOURCE Pan paniscus (isolate:bonobo-05) DNA.  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.  
1 (bases 1 to 1570)  
Liu,Y. and Saitou,N.  
Silver Project  
Published Only in DataBase (2000)  
REFERENCE 2 (bases 1 to 1570)  
AUTHORS Liu,Y. and Saitou,N.  
TITLE Direct Submission  
JOURNAL Submitted (11-APR-2000) Naruya Saitou, National Institute of  
Genetics, Laboratory of Evolutionary Genetics, 1111 Yata, Mishima,  
Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,  
URL:http://sayer.lab.nig.ac.jp/~silver/, Tel:81-559-81-6790,  
Fax:81-559-81-6789)  
FEATURES  
source  
1..1570  
/organism="Pan paniscus"  
/isolate="bonobo-05"  
/db\_xref="taxon:9597"  
/note="human sequence used for primer design based on Acc#  
AB020337"  
exon <1..20  
/number=3  
intron 21..667  
/number=3  
gene 668..1570  
/gene="betal,3-Galt 5"  
CDS 668..>1570  
/gene="betal,3-Galt 5"  
/codon\_start=1  
/product="UDP-Gal:GlcNAC betal,3-galactosyltransferase 5"  
/protein\_id="BAA94500.1"  
/db\_xref="GI:7593025"  
/translation="MAFFKMLRMVYCLLVGLALCLVFSMLNLFKEQSFVYKKGNF  
LKLPDTCROTPEPLVLVLTSSHOLAERMAQLVQWTKERTVKGKQLTFELGTTSS  
AAEKEVDQESQRHGDITQKDLGDFYNNLTLMGIEWHVRFCPOAFAFYKMDSDMF  
INVDYLLKLNKNTFFFTFFGLKNEFPKPFKWFVSKSEYPMWDRYPFPCSGVG  
YVFGSDVASQVYNVSESVYIKLEDVFGVGLERLNRLEELHSQPTFFPGGLRFSVC  
RFRIRVACHFKPRTLLDYQALEN"  
exon 668..>1570  
/gene="betal,3-Galt 5"  
/number=4  
BASE COUNT 408 a 372 c 392 g 398 t  
ORIGIN  
Query Match 94.9%; Score 797; DB 9; Length 1570;  
Best Local Similarity 98.8%; Pred. No. 3.8e-203;  
Matches 803; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
Qy 1 TTCAAAGAACAGTCTTGTGTTTACAAAGAGACGGGAATCTCTTAAGCTCCCAATACA 60  
|||||  
Db 758 TTCAAAGAACAGTCTTGTGTTTACAAAGAGATGGGAATCTCTTAAGCTCCCAATACA 817  
|||||

Qy 61 GACTGCAGGCGAGACACCTCCCTTCCTCGTCTGCTGGTACCTCATCCCAACAACAGTTG 120  
|||||  
Db 818 GACTGCAGGCGAGACACCTCCCTTCCTCGTCTGCTGGTACCTCATCCCAACAACAGTTG 877  
|||||  
Qy 121 GCTGAGCGCATGGCCATCCGCGAGAGCTGGGGGAAAGAGAGGATGTTGAAGGAAAGCAG 180  
|||||  
Db 878 GCTGAGCGCATGGCCATCCGCGAGAGCTGGGGGAAAGAGAGGACGGTGAAGGAAAGCAG 937  
|||||  
Qy 181 CTGAAGACATCTTCTCTCTGGGACCAACAGCAGTGCAGCGGAAAGAGAGGTGGAC 240  
|||||  
Db 938 CTGAAGACATCTTCTCTCTGGGACCAACAGCAGTGCAGCGGAAAGAGAGGTGGAC 997  
|||||  
Qy 241 CAGGAGAGCCAGCGACACCGGGACATTTCCAGAGAGGATTTCTAGACGCTCTATTACAAT 300  
|||||  
Db 998 CAGGAGAGCCAGCGACACCGGGACATTTCCAGAGAGGATTTCTAGACGCTCTATTACAAT 1057  
|||||  
Qy 301 CTGACCCTGAAGACCATGATGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGCGCGCG 360  
|||||  
Db 1058 CTGACCCTGAAGACCATGATGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGCGCGCG 1117  
|||||  
Qy 361 TTTGTGATGAAGACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGACTGCTT 420  
|||||  
Db 1118 TTTGTGATGAAGACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGACTGCTT 1177  
|||||  
Qy 421 CTGAAGAAAACAGACAAACACAGGTTTTCCTGCTGCTTCTTGAACCTCAATGAGTTTCCC 480  
|||||  
Db 1178 CTGAAGAAAACAGACAAACACAGGTTTTCCTGCTGCTTCTTGAACCTCAATGAGTTTCCC 1237  
|||||  
Qy 481 ATCAGGCGAGCATTCAGCAAGTGGTTTTCAGTAATCTGAATATCCGTGGGACAGGTAC 540  
|||||  
Db 1238 ATCAGGCGAGCATTCAGTAAGTGGTTTTCAGTAATCTGAATATCCGTGGGACAGGTAC 1297  
|||||  
Qy 541 CCACCATTCGCTCCGCGACCGGCTACGTTTCTGCGGACGTCGCGAGTTCAGGTGTAC 600  
|||||  
Db 1298 CCACCATTCGCTCCGCGACCGGCTACGTTTCTGCGGACGTCGCGAGTTCAGGTGTAC 1357  
|||||  
Qy 601 AATGCTCCAGAGCGTCCCATACATTAATACTGAGAGAGCTGTTTGTGGGCTCTGCCTC 660  
|||||  
Db 1358 AATGCTCCAGAGCGTCCCATACATTAATACTGAGAGAGCTGTTTGTGGGCTCTGCCTC 1417  
|||||  
Qy 661 GAAAGGCTGAACATCAGATTGGAGGAGTCCACTCCACGCGGACCTTTTCCAGGGGGC 720  
|||||  
Db 1418 GAAAGGCTGAACATCAGATTGGAGGAGTCCACTCCACGCGGACCTTTTCCAGGGGGC 1477  
|||||  
Qy 721 TTACGCTTCCTCGTATGCTCTTTCAGGAGGATCGTGCCCTGCCCTCATCAAGCCTCGG 780  
|||||  
Db 1478 TTACGCTTCCTCGTATGCTCTTTCAGGAGGATCGTGCCCTGCCCTCATCAAGCCTCGG 1537  
|||||  
Qy 781 ACTCTCTTGACTACTGGCAGGCTCTAGAGAA 813  
|||||  
Db 1538 ACTCTCTTGACTACTGGCAGGCTCTAGAGAA 1570  
|||||  
RESULT 10  
AB041413 1565 bp DNA linear PRI 13-APR-2000  
LOCUS Homo sapiens betal,3-Galt 5 gene for UDP-Gal:GlcNAC  
DEFINITION betal,3-galactosyltransferase 5, partial cds.  
ACCESSION AB041413  
VERSION AB041413.1 GI:7593020  
KEYWORDS  
SOURCE Homo sapiens (isolate:human-NR) DNA.  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1565)  
Liu,Y. and Saitou,N.  
Silver Project  
Published Only in DataBase (2000)  
REFERENCE 2 (bases 1 to 1565)  
AUTHORS Liu,Y. and Saitou,N.  
TITLE Direct Submission  
JOURNAL Submitted (11-APR-2000) Naruya Saitou, National Institute of

Genetics, Laboratory of Evolutionary Genetics, 1111 Yata, Mishima,  
Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,  
URL: <http://sayer.lab.nig.ac.jp/~silver/>, Tel:81-559-81-6790,  
Fax:81-559-81-6789)

FEATURES	source	Location/Qualifiers
exon		Location/Qualifiers
intron		Location/Qualifiers
gene		Location/Qualifiers
CDS		Location/Qualifiers
exon		Location/Qualifiers
BASE COUNT	410 a	370 c 385 g 398 t 2 others
ORIGIN		
Query Match	94.0%	Score 789.2; DB 9; Length 1565;
Best Local Similarity	99.8%	Pred. No. 4.7e-201;
Matches	799; Conservative	2; Mismatches 0; Indels 1; Gaps 1;
QY 1	TTCAAGAACAGTCCTTTGTTTACAGAAAGACGGGAACTTCCTTAAGCTCCCAAGATACA	60
Db	765 TTCAAGAACAGTCCTTTGTTTACAGAAAGACGGGAACTTCCTTAAGCTCCCAAGATACA	824
QY 61	GACTGCAGGACAGACCTCCCTTCCTCGTCTGCTGTGACCTCATCCCAACACAGTTG	120
Db	825 GACTGCAGGACAGACCTCCCTTCCTCGTCTGCTGTGACCTCATCCCAACACAGTTG	884
QY 121	GCTGAGCGCATGGCCATCCCGCAGACGTGGGGGAAACAGAGGATGGTGAAGGGAAAGCAG	180
Db	885 GCTGAGCGCATGGCCATCCCGCAGACGTGGGGGAAACAGAGGATGGTGAAGGGAAAGCAG	944
QY 181	CTGAAGACATTTCTCTCTGGGGACACACAGCAGTGCAGCGGAAACAGAAAGAGTGGAC	240
Db	945 CTGAAGACATTTCTCTCTGGGGACACACAGCAGTGCAGCGGAAACAGAAAGAGTGGAC	1004
QY 241	CAGGAGACCCAGGACACACGGGGACATTTATCCAGAAGATTTCTAGACGTCATTAACAAT	300
Db	1005 CAGGAGACCCAGGACACACGGGGACATTTATCCAGAAGATTTCTAGACGTCATTAACAAT	1064
QY 301	CTGACCTGAAGACCATGATGGGCATAGAATGGGTCCATCGCTTTTCTCTCGAGCGGCG	360
Db	1065 CTGACCTGAAGACCATGATGGGCATAGAATGGGTCCATCGCTTTTCTCTCGAGCGGCG	1124
QY 361	TTGTGATGAACACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAAGTCCCT	420
Db	1125 TTGTGATGAACACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAAGTCCCT	1184
QY 421	CTGAAGAAAAACAGAACACACAGGTTTTTTCACGTGGCTTCTTGAAACCTCAATGAGTTTCCC	480
Db	1185 CTGAAGAAAAACAGAACACACAGGTTTTTTCACGTGGCTTCTTGAAACCTCAATGAGTTTCCC	1244
QY 481	ATCAGGAGCCATTCAGCAAGTGGTTGTGTCAGTAAATCTGAAATATCCGTGGGACAGTAC	540



BASE COUNT		404 a		374 c		394 g		393 t		1 others	
ORIGIN											
/gene="betal.3-galt 5"											
Query Match		93.5%; Score 785.6; DB 9; Length 1566;									
Best Local Similarity		98.68; Pred. No. 4.4e-200;									
Matches 791; Conservative		1; Mismatches 10; Indels 0; Gaps 0;									
Qy	1	TTCAAGAACAGTCCCTTTGTTT	TTCAAGAACACGGGAAC	TCTCCTT	TAAGCTCC	CAGATACA	60				
Db	765	TTCAAGAACAGTCCCTTTGTTT	TTCAAGAACACGGGAAC	TCTCCTT	TAAGCTCC	CAGATACA	824				
Qy	61	GACTGCAGGACACACCTCCCTT	CTCTCGTCTG	TGGTGA	CTCATC	CCCAACAGTTG	120				
Db	825	GACTGCAGGACACACCTCCCTT	CTCTCGTCTG	TGGTGA	CTCATC	CCCAACAGTTG	884				
Qy	121	GCTGAGCGATGCCATCCG	CACGNGGGGA	AGAGAGAT	GCTGA	AGGGAACGAG	180				
Db	885	GCTGAGCGATGCCATCCG	CACGNGGGGA	AGAGAGAT	GCTGA	AGGGAACGAG	944				
Qy	181	CTGAAGACATCTTCCTCCT	GGGACCAC	CAGCAGT	GCAGCG	AAACGAAAGAGTGGAC	240				
Db	945	CTGAAGACATCTTCCTCCT	GGGACCAC	CAGCAGT	GCAGCG	AAACGAAAGAGTGGAC	1004				
Qy	241	CAGGAGAGCCAGGACAC	GGGGACAT	TATCCAGA	AGGATTT	CCTAGACGTCATTACAAT	300				
Db	1005	CAGGAGAGCCAGGACAC	GGGGACAT	TATCCAGA	AGGATTT	CCTAGACGTCATTACAAT	1064				
Qy	301	CTGACCCTGAAGACCAT	GATGGG	CATAGAA	TGGTCC	ATCGCTTTTGCTCAGCGCGG	360				
Db	1065	CTGACCCTGAAGACCAT	GATGGG	CATAGAA	TGGTCC	ATCGCTTTTGCTCAGCGCGG	1124				
Qy	361	TTTGTGATGAAACAGAC	TCTCAGACAT	GTTCAT	CAATGTT	GACTATCTGACTGAACTGCTT	420				
Db	1125	TTTGTGATGAAACAGAC	TCTCAGACAT	GTTCAT	CAATGTT	GACTATCTGACTGAACTGCTT	1184				
Qy	421	CTGAAGAAAAACAGAA	CAACAGGTTT	TTCAC	TGGCTTCT	TGAAACTCAATGAGTTTCCC	480				
Db	1185	CTGAAGAAAAACAGAA	CAACAGGTTT	TTCAC	TGGCTTCT	TGAAACTCAATGAGTTTCCC	1244				
Qy	481	ATCAGGACGCCATT	CAGCAAGT	GTGTTG	TCAAGTAA	TCTGAATATCCGTGGGACAGGTAC	540				
Db	1245	ATCAGGACGCCATT	CAGTAA	GTGTTG	TCAAGTAA	TCTGAATATCCGTGGGACAGGTAC	1304				
Qy	541	CCACCATTCGTCCG	CACCGGTAC	GTGTTT	CTGGCAG	CTGCGAGTCAGGTGTAC	600				
Db	1305	CCGCATTCGTCCG	CACCGGTAC	GTGTTT	CTGGCAG	CTGCGAGTCAGGTGTAC	1364				
Qy	601	AATGCTCCAGAGCGT	CCCATACAT	TAAACT	TGGAAG	ACGCTTTTGTGGGCTCTGGCTC	660				
Db	1365	AATGCTCCAGAGCGT	CCCATACAT	TAAACT	TGGAAG	ACGCTTTTGTGGGCTCTGGCTC	1424				
Qy	661	GAAGGCTGAACAT	CATAGATT	GGAGAG	CTCCACT	CCCGACCGACCTTTTTTCCAGGGGC	720				
Db	1425	GAAGGCTGAACAT	CATAGATT	GGAGAG	CTCCACT	CCCGACCGACCTTTTTTCCAGGGGC	1484				
Qy	721	TTACGCTTCCGTAT	GCCCTTTC	CAGAGG	ATCTG	GGCCTGCCACTTTCATCAAGCCTCG	780				
Db	1485	TTACGCTTCCGTAT	GCCCTTTC	CAGAGG	ATCTG	GGCCTGCCACTTTCATCAAGCCTCG	1544				
Qy	781	ACTCTCTTGGACT	TACTGG	CAGG	802						
Db	1545	ACTCTCTTGGACT	TACTGG	CAGG	1566						
RESULT 12											
AB041412											
LOCUS	AB041412	1360 bp	DNA	linear	PRI 17-OCT-2000						
DEFINITION	Gorilla gorilla betal.3-galt 5 gene for UDP-Gal:GlcNAc										
AB041412	betal.3-galactosyltransferase 5, partial cds.										
ACCESSION	AB041412										
VERSION	AB041412.1 GI:7593018										



```
|||||
Db 1544 CGGACTCTCTGGACTACTGGCAGGCTCTAGAGAT 1579
|||||
RESULT 14
AF254738 927 bp DNA linear ROD 24-JUL-2000
LOCUS Mus musculus beta 1,3-galactosyltransferase-V (B3gt5) gene,
DEFINITION complete cds.
ACCESSION AF254738
VERSION AF254738.1 GI:9246994
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 927)
Zhou,D., Henion,T.R., Jungalwala,F.B., Berger,E.G. and Hennet,T.
The beta 1,3-galactosyltransferase beta 3galnt-V is a stage-specific
embryonic antigen-3 (SSEA-3) synthase
J. Biol. Chem. 275 (30), 22631-22634 (2000)
20390006
PUBMED 10837462
REFERENCE 2 (bases 1 to 927)
Zhou,D. and Hennet,T.
Direct Submission
Submitted (11-APR-2000) Physiology, University of Zurich,
Winterthurerstrasse 190, Zurich 8057, Switzerland
JOURNAL Location/Qualifiers
FEATURES
source
1..927
/organism="Mus musculus"
/strain="129/SvJ"
/db_xref="taxon:10090"
<1..>927
/gene="B3gt5"
<1..>927
/gene="B3gt5"
/product="beta 1,3-galactosyltransferase-V"
1..927
/gene="B3gt5"
/codon_start=1
/product="beta 1,3-galactosyltransferase-V"
/protein_id="AAF86241.1"
/db_xref="GI:9246995"
/translation="MAHMKRLVYASILMMGALCLYFSMDSFRELFPVFKSHGKFLQ
IPDIDKQKPPFLVLLVTSYSHKQLAARIMRTWGRVTSVQGVQVTFLLGTSDSDE
EMDATTLESEQRDLIIQDKDAYENLTLKTMGMWVYHFCPTAYVMKTDSDMFVN
VGYLTELKKNKTRTFGTGKPHDFPIQKFNKFWKSEYVWPDRYPFCGCTGVV
FSSDAIVQVNVSSVPPKILEDFVGLCLAKLIRPELHTKQTFPPGLRFSVCRF
QKIVACHFMKRPQDLTYWQALENSKEQDCPAV"
BASE COUNT 226 a 241 c 235 g 225 t
ORIGIN
Query Match 59.8%; Score 502.2; DB 10; Length 927;
Best Local Similarity 75.6%; Pred. No. 6.4e-124;
Matches 637; Conservative 0; Mismatches 203; Indels 3; Gaps 1;
QY 1 TTCAAAGAACAGTCTTGGTTTACAGAAA---GACGGGAACCTTCCTTAAGCTCCCAAGAT 57
Db 82 TTCAGAAACCTCCCGTTTGTATTTTAAAGAAAGTCCGGGAAGTTCCTTCAGATTCCOGAT 141
QY 58 ACAGACTGCAGGACAGACACCTCCCTTCCTCGCTGCTGGTGACCTCCACCAACAG 117
Db 142 ATAGACTGCAAGACAGAGCCGCCCTTCTCGTGGTCTGCTGCTGCTGCTGCTGCTGCTG 201
QY 118 TTGGCTGAGCGCATGGCCATCCGGCAGAGCTGGGGGAAAGAGAGGATGGTGAAGGGAAG 177
Db 202 CTGGCCCTCGCATGGCCATCCGCAAGAGCTGGGGTAGAGAGACATCTGTGCGAGGCCAA 261
QY 178 CAGCTGAAGACATTTCTCTCTCTGGGGACCACACAGCAGTCGACGCGGAAACAGAGAGTG 237
Db 262 CAGGTGAGGACCTTCTCTCTCTGGGGACCTCCGACAGACCAGGAGATGGAGCCACA 321
```



Db 103259 GGTTCAGCTTTCCGCTGCGCTTTCAGAAAATTGTGGCATGCCATTTATGAAGCC 103318  
QY 778 CGGACTCTCTTGGAAGTCTAGAGAAATCCCGGGGGAAGATTGTCCGCCT 837  
Db 103319 CAGGACCTGCTCACTTACTGGCAAGCACTGGAGAACTCGAAGAACAGGACTGTCCTGCT 103378  
QY 838 GTC 840  
Db 103379 GTC 103381

Search completed: April 11, 2003, 20:24:54  
Job time : 1986.4 secs

GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 14:50:38 ; Search time 139.586 Seconds  
(without alignments)  
13552.085 Million cell updates/sec

Title: US-09-914-152-2\_COPY\_492\_1331  
Perfect score: 840  
Sequence: 1 ttcaaagaacagctcttgtt.....gggaagattgtccgcgtgc 840

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	840	100.0	2775	21 AAA93875	Human beta-1,3 gal
2	840	100.0	10562	21 AAA93876	Human betaGal-T5
3	836.8	99.6	1011	21 AAA27959	Human betaGal-T5
4	163.4	19.5	1739	15 AAQ67067	Beta-1,3-galactosyl
5	130.4	15.5	1773	20 AAX87193	Human Dendriac cDN
6	130.4	15.5	1897	24 ABK51201	Human cDNA encodin
7	130.4	15.5	2095	21 AAZ65022	Membrane-bound pro
8	130.4	15.5	2095	22 AAS46005	Human DNA encoding
9	130.4	15.5	2095	22 AAF92075	Human PRO1074 cDNA

10	130.4	15.5	2095	22 AAF44168	Human PRO1074 (UNQ
11	130.4	15.5	2168	22 AAH15711	Human cDNA sequenc
12	130.4	15.5	2189	20 AAX87192	Human Dendriac cDN
13	122.4	14.6	1266	20 AAX35710	cDNA encoding a pr
14	122.4	14.6	2420	20 AAX35711	cDNA encoding a pr
15	114.8	13.7	1134	21 AAS58791	DNA encoding the b
16	108.2	12.9	1134	22 AAD09356	Human beta-1,3-gal
17	93.2	11.1	1725	23 ABL14319	Drosophila melanog
18	93.2	11.1	4218	23 ABL14318	Drosophila melanog
19	89.8	10.7	1116	19 AAV49599	Human epidermoid c
20	89.8	10.7	1208	24 AAS16945	Human beta1,3-N-ac
21	89.8	10.7	2180	22 AAF29257	Human beta 1,3-N-a
22	89.8	10.7	2186	19 AAV49598	Human epidermoid c
23	89.8	10.7	2186	24 ABL14194	Nucleotide sequenc
24	89.8	10.7	2198	22 AAF93831	Human cDNA encodin
25	89.8	10.7	2205	22 AAF29256	Human beta 1,3-N-a
26	89.8	10.7	2210	22 AAS46068	Human DNA encoding
27	89.8	10.7	2427	23 ABV22526	Human prostate exp
28	89.8	10.7	2427	23 ABV28341	Human prostate exp
29	87.6	10.4	1337	24 AAS16946	Murine beta1,3-N-a
30	83.4	9.9	3033	23 ABL14314	Drosophila melanog
31	83.4	9.9	3330	23 ABL14322	Drosophila melanog
32	81.8	9.7	1271	20 AAX97916	Human secreted pro
33	81.8	9.7	1271	20 AAX87194	Human Brainiac-3 c
34	78.4	9.3	1296	22 AAF29258	Human beta 1,3-N-a
35	78.4	9.3	1420	21 AAF29258	DNA encoding the b
36	78.4	9.3	1642	21 AAC58115	Human PRO4344 nucl
37	78.4	9.3	1643	24 ABK69978	cDNA encoding huma
38	77.8	9.3	1359	23 AAS84438	DNA encoding novel
39	77.8	9.3	1360	24 AAD35225	Human TRNFR-10 cDN
40	74.6	8.9	977	21 AAS57383	DNA encoding a Bra
41	74.6	8.9	2117	24 AAC35219	Human TRNFR-4 cDNA
42	74.6	8.9	2684	21 AAC58117	Human PRO4397 nucl
43	74.6	8.9	2684	24 ABK69977	cDNA encoding huma
44	70.2	8.4	400	21 AAZ94080	Haematopoietic ste
45	68.6	8.2	1191	21 AAS45023	Human beta-1,3-gal

ALIGNMENTS

RESULT 1  
AAA93875  
ID AAA93875 standard; DNA; 2775 BP.  
XX AAA93875;  
AC  
DT 15-JAN-2001 (first entry)  
XX  
DE Human beta-1,3 galactose transferase encoding DNA.  
KW Beta-1,3 galactose transferase; treatment; diagnosis; cancer; human;  
KW digestive system; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200050508-A1.  
XX  
PD 31-AUG-2000.  
XX  
PF 24-FEB-2000; 2000WO-JP01070.  
XX  
PR 25-FEB-1999; 99JP-0047571.  
XX  
PA (KYOW ) KYOWA HAKKO KOGYO KK.  
XX  
PI Narimatsu H, Isshiki S, Togayachi A, Sasaki K;  
XX WPI: 2000-549409/50.  
DR P-PSDB; AAB93875.  
XX  
PT Beta-1,3 galactose transferase and DNA encoding it, useful for  
synthesis of type 1 sialyl Lewis, a carbohydrate for treatment of

PT	digestive system cancer
XX	
PS	Claim 5; Page 99-102; 123pp; Japanese.
XX	
CC	This invention relates to a polypeptide (I) with beta-1,3 galactose
CC	transferase activity, or variants of (I) comprising amino acid additions,
CC	deletions and/or substitutions. Included in the invention is DNA encoding
CC	all or part of (I); expression vectors containing the DNA, host cells
CC	transformed by the vectors; a method for the preparation of the
CC	polypeptide by culture of the transformants or by expression in the milk
CC	of a transgenic mammal, and antibodies recognising (I). The Beta-1,3
CC	galactose transferase protein transfers galactose by beta-1,3 bonding to
CC	N-acetylglucosamine present in a non-cyclic carbohydrate chain (such as
CC	GlcNAc-beta1-3gal-beta1-4Glc) to give Galbeta1-3GlcNAc. The protein and
CC	DNA encoding it are useful for the treatment and diagnosis of cancer of
CC	the digestive system. The present sequence represents Beta-1,3 galactose
CC	transferase encoding DNA.
XX	
SQ	Sequence 2775 BP: 681 A; 698 C; 669 G; 727 T; 0 other;
	Query Match 100.0%; Score 840; DB 21; Length 2775;
	Best Local Similarity 100.0%; Pred. No. 8,3e-250;
	Matches 840; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy	1 TTCAAGAACAGTCCTTTGTGTACAGAAAGACGGGAATCCTTAAAGTCGCCAGATACA 60
Db	492 TTCAAGAACAGTCCTTTGTGTACAGAAAGACGGGAATCCTTAAAGTCGCCAGATACA 551
Qy	61 GACTCGAGGCAGACACCTCCCTTCTCTGTCCTGCTGGTGACCTCATCCCACAAACAGTTG 120
Db	552 GACTCGAGGCAGACACCTCCCTTCTCTGTCCTGCTGGTGACCTCATCCCACAAACAGTTG 611
Qy	121 GCTGAGCCATGGCCATCCGGCAGACGTGGGGAAAGAGAGGATCGTGAAGGAAAGCAG 180
Db	612 GCTGAGCCATGGCCATCCGGCAGACGTGGGGAAAGAGAGGATCGTGAAGGAAAGCAG 671
Qy	181 CTGAAGACATTCTTCCTCTGGGACCACACAGTAGTGACGGGAACAAAAGAGGTGGAC 240
Db	672 CTGAAGACATTCTTCCTCTGGGACCACACAGTAGTGACGGGAACAAAAGAGGTGGAC 731
Qy	241 CAGGAGCCAGCCAGCGGGGACATTATCCAGAGGATTTCTAGACGCTATTACAAT 300
Db	732 CAGGAGCCAGCCAGCGGGGACATTATCCAGAGGATTTCTAGACGCTATTACAAT 791
Qy	301 CTGACCTTGAAGACCATGATGGGATAGAATGGGTCCATCGCTTTGTCTCAGCGCGCG 360
Db	792 CTGACCTTGAAGACCATGATGGGATAGAATGGGTCCATCGCTTTGTCTCAGCGCGCG 851
Qy	:61 TTTGTGATGAACAGACTCAGACATGTCATCAATGTTGACTATCTGACTGAACGTGT 420
Db	852 TTTGTGATGAACAGACTCAGACATGTCATCAATGTTGACTATCTGACTGAACGTGT 911
Qy	421 CTGAAGAAAAACAGAACACACCGTTTTTCACTCGCTTCTTGAACCTCAATGAGTTTCCC 480
Db	912 CTGAAGAAAAACAGAACACACCGTTTTTCACTCGCTTCTTGAACCTCAATGAGTTTCCC 971
Qy	481 ATCAGGCAGCCATTACAGCACTGTTTGTCAATAAATCTGAATATCCGTGGGACAGGTAC 540
Db	972 ATCAGGCAGCCATTACAGCACTGTTTGTCAATAAATCTGAATATCCGTGGGACAGGTAC 1031
Qy	541 CCACCATCTGCTCCGGACCGGTACGTGTTTCTTGCGGACGTGGGAGTCAGGTGTAC 600
Db	1032 CCACCATCTGCTCCGGACCGGTACGTGTTTCTTGCGGACGTGGGAGTCAGGTGTAC 1091
Qy	601 AATGCTCCAAGACGGTCCCATACATTAAACTGAAGACGTGTTTGTGGGCTCTGCCTC 660
Db	1092 AATGCTCCAAGACGGTCCCATACATTAAACTGAAGACGTGTTTGTGGGCTCTGCCTC 1151
Qy	661 GAAGAGGTGAACATCAGATTGGAGGAGCTCCACTCCCGACCGACCTTTTTCAGGGGGC 720
Db	1152 GAAGAGGTGAACATCAGATTGGAGGAGCTCCACTCCCGACCGACCTTTTTCAGGGGGC 1211
Qy	721 TTAGCTTTCGGTATCGCTCTTTCAGGAGGATCGTGGCTGCCACTTCATCAAGCCCTCG 780

D	B	1212	TTAGCGTCTCCGATGCCTCTTCAGGAGGATCGTGCCCTGCCACTTCATCAAGCCTCGG	1271
Q	Y	781	ACTCTCTTGAGTACTGGCAGGCTCTAGAGAATTCCGGGGGGAAGATTTCGCGCTGTC	840
D	B	1272	ACTCTCTTGAGTACTGGCAGGCTCTAGAGAATTCCGGGGGGAAGATTTCGCGCTGTC	1331
RESULT 2				
A	A	AAA93876		
I	D	AAA93876 standard; DNA; 10562 BP.		
X	X	AAAC		
X	X	AAA93876;		
D	T	15-JAN-2001 (first entry)		
X	X	Human beta3Gal-T5 encoding DNA.		
D	E			
K	W	Beta-1,3 galactose transferase; treatment; diagnosis; cancer; human;		
K	W	digestive system; beta3Gal-T5; ds.		
X	X	Homo sapiens.		
O	S			
X	X	WO2000050608-A1.		
P	N			
X	X	31-AUG-2000.		
X	X			
P	F	24-FEB-2000; 2000WO-JP01070.		
X	X			
P	R	25-FEB-1999; 99JP-0047571.		
X	X			
P	A	(KYOW ) KYOWA HAKKO KOGYO KK.		
X	X			
P	I	Narimatsu H, Isshiki S, Togayachi A, Sasaki K;		
X	X			
D	R	WPI; 2000-549409/50.		
X	X			
P	T	Beta-1,3 galactose transferase and DNA encoding it, useful for		
P	T	synthesis of type 1 sialyl Lewis, a carbohydrate for treatment of		
P	T	digestive system cancer		
X	X			
P	S	Claim 31; Page 103-111; 123pp; Japanese.		
X	X			
C	C	This invention relates to a polypeptide (I) with beta-1,3 galactose		
C	C	transferase activity, or variants of (I) comprising amino acid addition-		
C	C	deletions and/or substitutions. Included in the invention is DNA encod-		
C	C	all or part of (I); expression vectors containing the DNA, host cells		
C	C	transformed by the vectors; a method for the preparation of the		
C	C	polypeptide by culture of the transformants or by expression in the mil-		
C	C	of a transgenic mammal, and antibodies recognising (I). The Beta-1,3		
C	C	galactose transferase protein transfers galactose by beta-1,3 bonding to		
C	C	N-acetylglucosamine present in a non-cyclic carbohydrate chain (such as		
C	C	GlcNAc(betal-3galbetail-4Glc) to give Galbetail-3GlcNAc. The protein and		
C	C	DNA encoding it are useful for the treatment and diagnosis of cancer o-		
C	C	the digestive system. The present sequence represents a Beta3gal-T5		
C	C	encoding DNA sequence.		
X	X			
S	Q	Sequence 10562 BP; 2610 A; 2415 C; 2574 G; 2963 T; 0 other:		

Query Match	100.0%	Score 840;	DB 21;	Length 10562;
Best Local Similarity	100.0%;	Pred. No. 1.7e-249;		
Matches 840;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	TTCAAAGAACAGTCTCTTTGTTTACAAGAAAGACGGGAACCTTCCTTAAGCTCCAGATACA	60	
Db	8324	TTCAAAGAACAGTCTCTTTGTTTACAAGAAAGACGGGAACCTTCCTTAAGCTCCAGATACA	8383	
Qy	61	GACTGCAGGCAGACACCTCCCTTCCTCGTCTGCTGAGCTCATCCACAAACAGTTG	120	
Db	8384	GACTGCAGGCAGACACCTCCCTTCCTCGTCTGCTGAGCTCATCCACAAACAGTTG	8443	
Qy	121	GCTGAGCGCATGGCCATCCCGCAGACCTGGGGGAAAGAGAGGATGGTGAAGGGAAAGCAG	180	





QY	421	CTGAAGAAAAACAGAACACCAGGTTTTCACCTGGCTTCTTGAACATCAATGAGTTTCCC	480
Db	589	CTGAAGAAAAACAGAACACCAGGTTTTCACCTGGCTTCTTGAACATCAATGAGTTTCCC	648
QY	481	ATCAGGCAGCCATTACGCAAGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTAC	540
Db	649	ATCAGGCAGCCATTACGCAAGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTAC	708
QY	541	CCACCATTTCTGCTCCGGCACCGGTACGTGTTTCTTGGCGACGTGGGAGTCAAGGTGTAC	600
Db	709	CCACCATTTCTGCTCCGGCACCGGTACGTGTTTCTTGGCGACGTGGGAGTCAAGGTGTAC	768
QY	601	AATGTCTCCACAGCGTCCCATACATTAAACTGGAAGACGTGTTTCTTGGGGCTTGCCTC	660
Db	769	AATGTCTCCACAGCGTCCCATACATTAAACTGGAAGACGTGTTTCTTGGGGCTTGCCTC	828
QY	661	GAAGGCTGAACATCAGATTGGAGGAGCTCCACTCCCGCCGACCTTTTTCAGGGGGC	720
Db	829	GAAGGCTGAACATCAGATTGGAGGAGCTCCACTCCCGCCGACCTTTTTCAGGGGGC	888
QY	721	TTACGCTTCTCGTATGCGTCTTCAGGAGAGTCGTGGGCTGCCACTTCATCAAGCCTCG	780
Db	889	TTACGCTTCTCGTATGCGTCTTCAGGAGAGTCGTGGGCTGCCACTTCATCAAGCCTCG	948
QY	781	ACTCTCTTGGACTACTGCGAGGCTCTAGAGAAATCCCGGGGGGAAAGATTGTCGCGCTGTC	840
Db	949	ACTCTCTTGGACTACTGCGAGGCTCTAGAGAAATCCCGGGGGGAAAGATTGTCGCGCTGTC	1008
RESULT	4		
AAQ67067			
ID	AAQ67067	standard; cDNA to mRNA; 1739 BP.	
XX	XX		
AC	AAQ67067;		
XX	XX		
DT	14-MAR-1995	(first entry)	
XX	XX		
DE	Beta-1,3-galactosyltransferase cDNA.		
XX	XX		
KW	Beta-1,3-galactosyltransferase; enzyme; saccharide chain; pAMoPRWML;		
KW	KJM-1 cells; ds.		
XX	XX		
OS	Homo sapiens.		
XX	XX		
FH	Key	Location/Qualifiers	
FT	CDS	676..1656	
FT	FT	/*tag= a	
FT	FT	/product= Beta-1,3-galactosyltransferase	
XX	XX		
PN	JP06181759-A.		

PD	05-JUL-1994.	
XX		
XX	16-DEC-1992;	92JP-0336436.
XX		
XX	16-DEC-1992;	92JP-0336436.
PR		
XX	(KYOW ) KYOWA HAKKO KOGYO KK.	
PA		
XX	WPI; 1994-251683/31.	
XX	P-PSDB; AAR57433.	
XX		
PT	Beta-galactosyl-transferase DNA and protein - useful for prodn. of	
PT	saccharide chains	
XX		
XX	Claim 3; Page 22-24; 47pp; Japanese.	
PS		
PS		
XX		
CC	This sequence encodes a beta-1,3-galactosyltransferase. This enzyme	
CC	can be used to produce physiologically active saccharide chains and	
CC	variants, and for improvement of saccharide chains bound to	
CC	physiologically active proteins. This cDNA represents a fragment	
CC	of the plasmid pAMCPRWMI which was cloned in KJM-1 cells.	
CC		

XX	Sequence	1739 BP;	473 A;	407 C;	422 G;	437 T;	0 other;
Qy	Query Match	19.5%;	Score 163.4;	DB 15;	Length 1739;		
Best Local Similarity	55.7%;	Pred. No. 6.8e-40;					
Matches 358;	Conservative	0;	Mismatches 276;	Indels	9;	Gaps	
Qy	38	ACTTCCTTAAGCTCCCGACATACAGACTGCGAGGCGAGACACTCCTCTCTCTCGTCTCGTCTGCTGG	97				
Db	866	AATTTCTTATCAACGAGGCCAATAAATGTGAGAAAAACATTCCTTTTCTTTGTTATCTCTCA	925				
Qy	98	TGACCTCATCCCAACAACAGTTGGCTGAGCGCATGGCCATCCGGCAGACAGTGGGGGGAAG	157				
Db	926	TCAGACCACCTCAACAAGGAATTTGATGCCCGTCAGGCAATTCAGAGACAGCTGGGGGATG	985				
Qy	158	AGAGGATGGTGAAGGGAAGCAGCTGAAGACATCTTCCTCTCTGGGACACACGACGAGT	217				
Db	986	AGAACAACTTTAAGGGGATCAAGATAGCCACCGCTTCCTCTCTGGGCAAGATGCTGATC	1045				
Qy	218	CAGCGGAACGAAGAGGTGGACCGAGGAGGCGAGCGACGCGGGACATTCACGAAGG	277				
Db	1046	CTGTTCTCAATCAGATGGTGGAGCAGAGGCCAAATCTCCATGATATCATCTCGTGGAGG	1105				
Qy	278	ATTTCCTAGACGCTATTACAACTGACCCCTGAAGACCATGATGGGCATAGAATGGGTCC	337				
Db	1106	ACTTATTGACTCTTACCATAACCTTACCCTCAAAACATTAATGGGATCAGATGGGTGG	1165				
Qy	338	ATCGTTTGTGCTCAGCGGGCGTTTGATGAAGAACAGACTCAGACATGTTTCATCAATG	397				
Db	1166	CCACTTTTGTTCAAAGCCAAATGTGTCATGAAGACAGACAGCGACATTTTGTAAACA	1225				
Qy	398	TTGACTATCTGACTCAACTGCTTCTGAA-----GAAAAACAGAACACACAGGTTTTTCA	451				
Db	1226	TGGCAATCTTATTATTAATTTACTTGAACCCCTCCACCAAGCCACGAAGGATATTTTA	1285				
Qy	452	CTGGCTCTTTGAAACTCAATGAGTTTCCCATCAGGCGAGCATTCAGCAAGTGGTTTGTCA	511				
Db	1286	CTGCGTATGTCA---TTAATGGAGACCCGATTCGGGATGTCGCCAGTAAGTGGTATATGC	1342				
Qy	512	GTAATCTGGAATATCCGTGGGACAGGTACCCACCATTCGTCTCCGGCACCGGTACGTGT	571				
Db	1343	CCAGGGATTTGTACCCAGACAGTAATACTACCCACCTTCTGTTCTGGGGACTGGCTACATCT	1402				
Qy	572	TTTCTGGGACGTGGCGAGTCAGGTGTACAATGTCTTCCAAGAGCGTCCCATACATTAAAC	631				
Db	1403	TTTCAGCCGATGTAGCTGAACCTCATTTACAAGACCTCACTCCACACAAGGCTGCTTCAAC	1462				
Qy	632	TGGAAGACGTGTTTGTGGGGCTCTCCCTCGAAGGCTGAACAT	674				
Db	1463	TTGAAGACGTATATGTGGGACTGTGCTCTCGAAGCTGGGCAT	1505				

RESULT	5
AAX87193	
ID	AAX87193 standard; cDNA; 1773 BP.
XX	
AAx87193;	
XX	
DT	(first entry)
DE	Human Dendriac cDNA.
XX	
KW	Dendriac; Brainiac-2; human; Notch; immune disorder; neurological disorder; diagnosis; therapy; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
CDS	Location/Qualifiers 20..980
FT	/tag= a
FT	
sig_peptide	20..94
FT	/tag= b
mat_peptide	95..977
FT	

FT XX /\*tag= c  
PN WO9931116-Al.  
XX 24-JUN-1999.  
XX 17-DEC-1998; 98WO-US27049.  
XX 17-NOV-1998; 98US-0108928.  
PR 18-DEC-1997; 97US-0068006.  
PR 12-MAR-1998; 98US-0077687.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Ebner R, Endress GA, Florence KA, Rosen CA, Ruben SM;  
PI Soppet DR, Yu G;  
XX  
DR WPI: 1999-430031/36.  
DR P-PSDB; AAY06461.  
XX  
PT Human proteins for treating and detecting immune and neurological  
PT disorders  
XX  
PS Claim 2; Page 119-120; 132pp; English.  
XX  
CC This is the nucleotide sequence of a cDNA clone which codes for  
CC human Dendriac (see AAY06461), also called Brainiac-2, a novel  
CC member of the Brainiac family. Dendriac cDNA (see also AAY87193) is  
CC deposited as ATCC 203056 and ATCC 209627. It was initially  
CC discovered in a dendritic cell cDNA library. Additional clones of  
CC the same gene were identified in NTERA2 cell, adult pulmonary  
CC tissue, salivary gland, ovary, Caco-2 colon adenocarcinoma, smooth  
CC muscle, cerebellum, 8-week-old human myob, haemangiopericytoma,  
CC amygdala, substantia nigra and whole brain cDNA libraries. Dendriac  
CC nucleic acids are used in the recombinant production of Dendriac  
CC polypeptides, especially the mature protein or epitope-bearing  
CC fragments. Dendriac and Brainiac proteins are members of the Notch  
CC family of proteins involved in the control of cell division. The  
CC proteins may be used to detect and/or treat immune system and  
CC neurological disorders.  
XX  
SQ Sequence 1773 BP; 568 A; 300 C; 339 G; 566 T; 0 other;

Query Match 15.5%; Score 130.4; DB 20; Length 1773;  
Best Local Similarity 48.9%; Pred. NO. 1.2e-29;  
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;

QY 20 TTACAGAAAGACGGGAACCTTCTTAAGCTCCAGATACAGACTGCAGGAGACACCTC 79  
DB 157 TTTACAGACAAACATTTCACTTCACTTCGAGAGCAATTCAACTGCTCTCATCAAAATC 216  
QY 80 CTTCTCCTGCTGCTGTGAGCTCATCCCAACAAAGTTGGTGTGAGCGCATGCCATCC 139  
DB 217 CATTTCTGTCTATCTGTGACCTCCACCTCCAGATGTGAAGCCAGGAGCCATTA 276  
QY 140 GGCAGACGTGGGGGAAGAGAGATGGTGAAGGAAGAGAGCTGAAGACATCTTCCTCC 199  
DB 277 GAGTTACTTGGGGTGAAGAAAGATCTTGTGGGGATATGAGTTCTTACATTTTCTTAT 336  
QY 200 TGGGACACACAGCAGTGCAGCGGAACAAAGAGG-----TGGACAGGAGGCC 250  
DB 337 TAGGCCAGAGGCTGAAGGAAGCAAAATGTTGGCAATGCTCTTAGAGGATGAACACC 396  
QY 251 AGCGACACGGGACATATCCAGAAGGATTTCTTAGACGCTCTTACAAATCTGACCTGA 310  
DB 397 TTTCTTTATGGTGACATAATCCACAGATTTTGTAGACATATATAACTGACCTTGA 456  
QY 311 AGACCATGATGGCATAGATGGTCCATCGCTTTTGTCTCAGCGCGCTTTGTGATGA 370  
DB 457 AAACCAATTATGCAATTCAGTGGTGAACAGTGGTGGTGGTGGTGGTGGTGGTGGT 516  
QY 371 AAACAGACTCAGACATGTTCAATGTTGACTATCTGACTGAATGCTTCTGAAGAAA 430

DB 517 AGACAGACACTGATGTTTTCATCAATACTGGCAATTTAGTGAAGATATCTTTTAAACCTAA 576  
QY 431 ACAGAAACACACAGGTTTTTCTACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGCAGC 490  
DB 577 ACCACTCAGAGAGTGTTCACAGGTTATCTCTAATGATAATATTCCTATAGAGAT 636  
QY 491 CATTACAGCAAGTGTGTTGTCTAGTAAATCTGAATATCGTGGGACAGGTACCCACCATCT 550  
DB 637 TTTACCAAAAACCCATATTTCTTACCAGAGATATCTTTCAAGGTGTTCCTCCCTACT 596  
QY 551 GCTCCGGCACCGGTACGTGTTTCTGGGACGTGCGGAGTCAAGGTGTACAAATGTCTCCA 610  
DB 697 GCAGTGGGTGGTTATATAATGTCAGAGATTTGGTGCCCAAGGATCTATCAATGATGG 756  
QY 611 AGAGCTCCCATACATTAAACTGGAGAGAGCTGTTTGTGGGGCTCTGCTCGAAAGGCTGA 670  
DB 757 GTCAGTAAACCCCATCAAGTTTGAAGATGTTTATGTCGGGATCTGTTGAATTTATTA 816  
QY 671 ACATCAGATTGGAGGAGCTCCACTCCAGCCGACCTTTTTCAGGGGGCTTACGCTTCT 730  
DB 817 AAGTGAACATTCATATTCAGAGACACAAATCTTTTCTTCTATATAGAATCCATTTGG 876  
QY 731 CCGTATGCTCTTTCAGGAGGATCGTGGCTGCCACTTCATCAAGCCTCGGACTCTCTTGG 790  
DB 877 ATGCTGTCAACTGAGACGTGTGATTGCAGCCCATGCTTTTCTTCCAGGAGATCATCA 936  
QY 791 ACTACTGGCAGG 802  
DB 937 CTTTGTGGCAGG 948

RESULT 6  
ABK51201  
ID ABK51201 standard; cDNA; 1897 BP.  
XX  
AC ABK51201;  
XX  
DT 30-JUL-2002 (first entry)  
XX  
DE Human cDNA encoding betal,3-acetyl galactosamine transferase.  
XX  
KW Human; ss: gene; betal,3-acetyl galactosamine transferase;  
KW GB4 sugar chain.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT misc\_feature /tag= a  
FT 109..1101 /note= "This region is specifically claimed in claim 1"  
FT CDS 109..1104  
FT /tag= b  
FT /product= "Betal,3-acetyl galactosamine transferase"  
FT polyA\_signal 1157..1162  
FT /tag= c  
XX  
PN JP2002085069-A.  
XX  
PD 26-MAR-2002.  
XX  
PF 08-SEP-2000; 2000JP-0273835.  
XX  
PR 08-SEP-2000; 2000JP-0273835.  
XX  
PA (SEG ) SEIKAGAKU KOGYO CO LTD.  
PA (FURU/) FURUKAWA K.  
XX  
DR WPI: 2002-378274/41.  
DR P-PSDB; AAU80224.  
XX  
PT Preparation of beta-1,3-acetyl galactosamine transferase -  
XX  
PS Claim 1; Page 11-13; 15pp; Japanese.

XX The invention relates to the preparation of beta-1,3-acetylgalactosamine  
CC (a Gb4 sugar chain molecule) comprising introducing a DNA appearing as  
CC ABK51201 (S1) or a DNA hybridizing with S1 or a base sequence  
CC complementary to the base sequence or part of these base sequences under  
CC a stringent condition into a cell and growing the cell to express beta-1,  
CC 3-acetylgalactosamine transferase and collecting it. Also included is the  
CC preparation of Gb4 sugar chain comprising contacting a polypeptide  
CC comprising the beta-1,3-acetylgalactosamine transferase protein appearing  
CC as AAU80224 (A) or a polypeptide consisting of an amino acid sequence in  
CC which at least one amino acid is replaced, deleted, inserted or  
CC transferred in the amino acid sequence (A) and having enzymatic activity  
CC transferring N-acetylgalactosamine residue from an N-acetylgalactosamine  
CC donor to the C3 site of the galactose residue in Gb3 sugar chain which is  
CC the receptor with an N-acetylgalactosamine donor and Gb3 sugar chain.  
CC The present sequence is the cDNA encoding beta-1,3-acetylgalactosamine  
CC transferase.  
XX

SQ Sequence 1897 BP; 596 A; 327 C; 366 G; 608 T; 0 other;

Query Match 15.5%; Score 130.4; DB:24; Length 1897;  
Best Local Similarity 48.9%; Pred. NO. 1.2e-29;  
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;

QY 20 TTACAGAGAAAGCGGAACTTCCTTAAGCTCCAGATACAGACTGCAGGCGAGACACCTC 79  
DB 281 TTTACAGACAGACTTCTACTTCACACTTCGAGAGCAATCAAACTGCTCTCATCAAAATC 340  
QY 80 CTTCTCCTCGTCTGCTGTGACCTATCCCAACAAAGTTGGCTGAGCGCATGCCATCC 139  
DB 341 CATTTCTGTCTATCTGCTGTGACCTCCACCCTTCAGATGTGAAGCCAGGCGACCCATTA 400  
QY 140 GGCAGACGTGGGGAAGAGAGATGGTGAAGGAAGAGCTGAAGACATCTTCCTCC 199  
DB 401 GAGTTACTTGGGGTCAAAAAAGTCTGTGGGGATATGAGTTCTTACATTTTCTTAT 460  
QY 200 TGGGACACCAGCAGCTGCAGCGGAACAAAGAGG-----TGGACAGGAGAGCC 250  
DB 461 TAGGCCAAGAGCTGAAAGGAAGACAAATGTGGCAATTCCTTAGAGGATGAACACC 520  
QY 251 AGCGACACGGGACATATTCAGAGAGATTTCTAGACGCTCTATTACAATCTGACCCCTGA 310  
DB 521 TTCCTTATGGTGACATAATCCGACAGATTTTATAGACACATATATACTGACCTTGA 580  
QY 311 AGACATGATGGGCATAGATGGTCCATCGGTTTGTCTCAGGCGCGTTTGTGATGA 370  
DB 581 AAACCATTTATGGCATTCAGGTGGGTAACTGAGTTTGGCCCAATGCCAAGTACGTAATGA 640  
QY 371 AAACAGACTCAGACATGTTCAATGTTGACTATCTGACTGACTGACTCTCTGAAGAAA 430  
DB 641 AGACAGACACTGATGTTTCATCAATCTACTGGCAATTTAGTGAAGTATCTTTTAAACCTAA 700  
QY 431 ACAGAACACACGAGTTTTCCTGCTGCTTCTTGAAACTCAATGAGTTTCCCATCAGGCGAGC 490  
DB 701 ACCACTCAGAGAGTTTTCACAGGTTATCTCTTAATGATAATTAATCTATAGAGGAT 760  
QY 491 CATTCAGCAAGTGGTTTGTGCTAAATCTGAATATCCGTGGGACAGGTACCCACCTTCT 550  
DB 761 TTTACCAAAAACCCATATTTCTACAGAGATATCTTCAAGGTGTTCCCTCCATCTACT 820  
QY 551 GCTCCGCGACCGGTACGTGTTTCTGCGAGCTGCGGAGTCAAGTGTACAATGCTCTCCA 610  
DB 821 GCAGTGGGTTGGTTATATAATGTCCAGAGATTTGTTGCCAAGGATCTATGAATGATGG 880  
QY 611 AGACGGTCCCATACATTAATAACTGGAAGAGCTGTTTGTGGGCTCTGCTCGAAAGGCTGA 670  
DB 881 GTCACGTAAAACCCATCAAGTTTGAAGATGTTTATTCGGGATCTCTTTGAATTTATTA 940  
QY 671 ACATCAGATTGGAGAGCTCCACTCCAGCCGACCTTTTTCAGGGGCTTACGCTTCT 730  
DB 941 AAGTGAACATTCATATCCAGAGACACAAATCTTTTCTTCTATATAGAATCCATTTGG 1000  
QY 731 CCGTATGCCTCTTCAGGAGGATCGTGGCTGCCACTTTCATCAAGCCTCGGACTCTCTGG 790

DB 1001 ATGTCTCTCAACTGAGAGCTGTGATGGAGCCCATGGCTTTCTTCCAGGAGATCATCA 1060  
QY 791 ACTACTGGCAGG 802  
DB 1061 CTTTTTGGCAGG 1072  
RESULT 7  
AAZ65022  
ID AAZ65022 standard; cDNA: 2095 BP.  
XX AAZ65022;  
AC AAZ65022;  
XX 05-APR-2000 (first entry)  
XX Membrane-bound protein PRO1074 encoding cDNA.  
XX  
KW Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;  
KW Pharmaceutical; receptor immunoadhesin; gene mapping; ss.  
XX Homo sapiens.  
XX WO9963088-A2.  
XX 09-DEC-1999.  
XX 02-JUN-1999; 99WO-US12252.  
XX 02-JUN-1998; 98US-0087607.  
PR 02-JUN-1998; 98US-0087609.  
PR 02-JUN-1998; 98US-0087759.  
PR 03-JUN-1998; 98US-0087827.  
PR 04-JUN-1998; 98US-0088021.  
PR 04-JUN-1998; 98US-0088025.  
PR 04-JUN-1998; 98US-0088028.  
PR 04-JUN-1998; 98US-0088029.  
PR 04-JUN-1998; 98US-0088030.  
PR 04-JUN-1998; 98US-0088033.  
PR 04-JUN-1998; 98US-0088326.  
PR 05-JUN-1998; 98US-0088167.  
PR 05-JUN-1998; 98US-0088202.  
PR 05-JUN-1998; 98US-0088212.  
PR 05-JUN-1998; 98US-0088217.  
PR 09-JUN-1998; 98US-0088655.  
PR 10-JUN-1998; 98US-0088722.  
PR 10-JUN-1998; 98US-0088730.  
PR 10-JUN-1998; 98US-0088734.  
PR 10-JUN-1998; 98US-0088738.  
PR 10-JUN-1998; 98US-0088740.  
PR 10-JUN-1998; 98US-0088741.  
PR 10-JUN-1998; 98US-0088742.  
PR 10-JUN-1998; 98US-0088810.  
PR 10-JUN-1998; 98US-0088811.  
PR 10-JUN-1998; 98US-0088824.  
PR 10-JUN-1998; 98US-0088825.  
PR 10-JUN-1998; 98US-0088826.  
PR 11-JUN-1998; 98US-0088858.  
PR 11-JUN-1998; 98US-0088861.  
PR 11-JUN-1998; 98US-0088863.  
PR 11-JUN-1998; 98US-0088876.  
PR 12-JUN-1998; 98US-0089090.  
PR 12-JUN-1998; 98US-0089105.  
PR 16-JUN-1998; 98US-0089440.  
PR 16-JUN-1998; 98US-0089512.  
PR 16-JUN-1998; 98US-0089514.  
PR 17-JUN-1998; 98US-0089538.  
PR 17-JUN-1998; 98US-0089539.  
PR 17-JUN-1998; 98US-0089598.  
PR 17-JUN-1998; 98US-0089600.  
PR 17-JUN-1998; 98US-0089653.  
PR 18-JUN-1998; 98US-0089801.

```
PR 18-JUN-1998; 98US-0089907.
PR 18-JUN-1998; 98US-0089908.
PR 19-JUN-1998; 98US-0089947.
PR 19-JUN-1998; 98US-0089948.
PR 19-JUN-1998; 98US-0089952.
PR 22-JUN-1998; 98US-0090246.
PR 22-JUN-1998; 98US-0090252.
PR 22-JUN-1998; 98US-0090254.
PR 23-JUN-1998; 98US-0090349.
PR 23-JUN-1998; 98US-0090355.
PR 23-JUN-1998; 98US-0090429.
PR 24-JUN-1998; 98US-0090431.
PR 24-JUN-1998; 98US-0090435.
PR 24-JUN-1998; 98US-0090444.
PR 24-JUN-1998; 98US-0090445.
PR 24-JUN-1998; 98US-0090461.
PR 24-JUN-1998; 98US-0090472.
PR 24-JUN-1998; 98US-0090535.
PR 24-JUN-1998; 98US-0090538.
PR 24-JUN-1998; 98US-0090540.
PR 24-JUN-1998; 98US-0090557.
PR 25-JUN-1998; 98US-0090676.
PR 25-JUN-1998; 98US-0090678.
PR 25-JUN-1998; 98US-0090688.
PR 25-JUN-1998; 98US-0090690.
PR 25-JUN-1998; 98US-0090691.
PR 25-JUN-1998; 98US-0090694.
PR 25-JUN-1998; 98US-0090695.
PR 25-JUN-1998; 98US-0090696.
PR 26-JUN-1998; 98US-0090862.
PR 26-JUN-1998; 98US-0090863.
PR 01-JUL-1998; 98US-0091358.
PR 01-JUL-1998; 98US-0091360.
PR 01-JUL-1998; 98US-0091544.
PR 02-JUL-1998; 98US-0091478.
PR 02-JUL-1998; 98US-0091486.
PR 02-JUL-1998; 98US-0091519.
PR 02-JUL-1998; 98US-0091626.
PR 02-JUL-1998; 98US-0091628.
PR 02-JUL-1998; 98US-0091633.
PR 02-JUL-1998; 98US-0091646.
PR 02-JUL-1998; 98US-0091673.
PR 07-JUL-1998; 98US-0091678.
PR 07-JUL-1998; 98US-0091682.
PR 09-JUL-1998; 98US-0092182.
PR 10-JUL-1998; 98US-0092472.
PR 20-JUL-1998; 98US-0093339.
PR 30-JUL-1998; 98US-0094651.
PR 04-AUG-1998; 98US-0095282.
PR 04-AUG-1998; 98US-0095285.
PR 04-AUG-1998; 98US-0095301.
PR 04-AUG-1998; 98US-0095302.
PR 04-AUG-1998; 98US-0095318.
PR 04-AUG-1998; 98US-0095321.
PR 04-AUG-1998; 98US-0095325.
PR 10-AUG-1998; 98US-0095916.
PR 10-AUG-1998; 98US-0095929.
PR 10-AUG-1998; 98US-0096012.
PR 11-AUG-1998; 98US-0096143.
PR 11-AUG-1998; 98US-0096146.
PR 12-AUG-1998; 98US-0096329.
PR 17-AUG-1998; 98US-0096757.
PR 17-AUG-1998; 98US-0096766.
PR 17-AUG-1998; 98US-0096768.
PR 17-AUG-1998; 98US-0096773.
PR 17-AUG-1998; 98US-0096791.
PR 17-AUG-1998; 98US-0096867.
PR 17-AUG-1998; 98US-0096891.
PR 17-AUG-1998; 98US-0096894.
PR 17-AUG-1998; 98US-0096895.
PR 17-AUG-1998; 98US-0096897.
PR 18-AUG-1998; 98US-0096949.
PR 18-AUG-1998; 98US-0096950.

PR 18-AUG-1998; 98US-0096959.
PR 18-AUG-1998; 98US-0096960.
PR 18-AUG-1998; 98US-0097022.
PR 19-AUG-1998; 98US-0097141.
PR 20-AUG-1998; 98US-0097218.
PR 24-AUG-1998; 98US-0097661.
PR 26-AUG-1998; 98US-0097951.
PR 26-AUG-1998; 98US-0097952.
PR 26-AUG-1998; 98US-0097954.
PR 26-AUG-1998; 98US-0097955.
PR 26-AUG-1998; 98US-0097971.
PR 26-AUG-1998; 98US-0097974.
PR 26-AUG-1998; 98US-0097978.
PR 26-AUG-1998; 98US-0097979.
PR 26-AUG-1998; 98US-0097986.
PR 26-AUG-1998; 98US-0098014.
PR 31-AUG-1998; 98US-0098525.
PR 16-SEP-1998; 98US-0100634.
PR 12-JAN-1999; 99US-0115565.
XX
XX (GETH ) GENENTECH INC.
PA
XX
XX Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;
PI Wood WI, Yuan J;
XX
XX WPI; 2000-072883/06.
DR P-PSDB; AAY66685.
XX
XX Membrane-bound proteins and related nucleotide sequences
PT
XX
XX Claim 2; Fig 136; 822pp; English.
XX
XX The invention provides membrane-bound PRO polypeptides and
CC polynucleotides encoding them. The PRO sequences of the invention were
CC identified based on extracellular domain homology screening. The PRO
CC sequences have homology with proteins including LDL receptors, TIE
CC ligands and various enzymes. The membrane-bound proteins and receptor
CC molecules are useful as pharmaceutical and diagnostic agents. Receptor
CC immunoadhesins, for instance, can be used as therapeutic agents to block
CC receptor-ligand interactions. The membrane-bound proteins can also be
CC employed for screening of potential peptide or small molecule inhibitors
CC of the relevant receptor/ligand interaction. The PRO encoding sequences
CC are useful as hybridization probes, in chromosome and gene mapping and in
CC the generation of antisense RNA and DNA. PRO nucleic acid sequences
CC will also be useful for the preparation of PRO polypeptides, especially
CC by recombinant techniques.
XX
XX Sequence 2095 BP; 629 A; 382 C; 418 G; 666 T; 0 other;
SQ
Query Match 15.5%; Score 130.4; DB 21; Length 2095;
Best Local Similarity 48.9%; Pred. No. 1.3e-29;
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;

QY 20 TTTCAGAGAGACGGGAACTTCCTTAAGCTCCAGCATACAGACTGCAGGCAGACACCTC 79
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 80 CCTTCCTCGTCTGCTGGTGACTCTATCCACAAAGAGTTGGTGAGCGCATGCCATCC 139
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 140 GGCAGACGTGGGGAAAGAGAGAGATGGTGAAGGGAAGCAGCTGAACACATCTTCTCTCC 199
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 200 TGGGACCCAGCAGCAGTGCAGCGGAAACGAAAGAGG-----TGCACGAGGAGCC 250
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 674 TAGCCAGAGAGCTGAAAGGAGAGACAAATGTTGCAATTGCTCTAGAGATGAACACC 733
QY 251 AGGCACAGCGGACATTTATCCAGAAGGATTTCTCTAGACGCTCTATTACAATCTGACCTGA 310
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 734 TTCCTTATGTTGACATAATCCGACAAAGATTTTATAGACACATATATAACCTGACCTGA 793
```

Qy	311	AGACCATGATGGGCATAGAAATGGGTCCATCGCTTTTGTGCTCAGGGGGCGTTTGTGATGA	370
Db	794	AAACCATTTATGSCATTCAGTGGTAACTGAGTTTTGCCCCAATGCCAAGTACGTAATGA	853
Qy	371	AAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACCTGCTTCTGAAGAAA	430
Db	854	AGACAGACACTGATGTTTTCATCAATACTGCGAAATTTAGTGAAGTATCTTTTAAACCTAA	913
Qy	431	ACAGAAACACAGGTTTTTTCAGTGCTTCTTGAAACTCAATGAGTTTCCCATCAGGCAGC	490
Db	914	ACCACTCAGAGAAGTTTTTTCAGGTTATCCCTTAATTGATAATTTCTCTATAGAGGAT	973
Qy	491	CATTTCAGCAAGTGTTTGTGCATTAATCTGAATACCGTGGGACAGAGTACCCACCACTTCT	550
Db	974	TTTACCAAAAACCCATATTTCTTACCAGAGTATCCTTTCAGGTGTTCCCTCCCACTACT	1033
Qy	551	GTCGCGCACGGCTACGTGTTTTCTGGGACGTGCGGAGTCAGGTGTACAAATGCTCTCCA	610
Db	1034	GCAGTGGGTTGGGTTATATAATGTCAGAGATTGGTGCCAAGGATCTATGAATGATGG	1093
Qy	611	AGAGCGTCCCATACATTAACTGGAAGACCGTGTGTGGGGCTCTCGCTCGAAAGCGTGA	670
Db	1094	GTCAGCTAAAACCCATCAAGTTGAAGATGTTTATGTCGGGATCTTTTGAATTTATTAA	1153
Qy	671	ACATCAGATTGGAGGAGTCCACTCCCAGCCGACCTTTTTTCCAGGGGCTTACGCTTCT	730
Db	1154	AAGTGAACATTCAATTCCAGAAGACACAAATCTTTCTTTCTATATAGAATCCATTGG	1213
Qy	731	CCGTATGCCTCTTCAGGAGGATCGTGGCTGCGACTTCATCAAGCCTCGGACTCTCTTGG	790
Db	1214	ATGCTGTCAACTGAGACGTGTGATTGACGCCCATGGCTTTTCTTCCAGGAGATCATCA	1273
Qy	791	ACTACTGGCAGG	802
Db	1274	CTTTTGGCAGG	1285

RESULT A

RESULT B  
AAS46005  
ID AAS46005 standard; CDNA; 2095 BP.

AX  
AC

AA  
DT 18-DEC-2001 (first entry)

XX  
DE Human DNA encoding PRO polypeptide sequence #81.

KW pro polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep; ss;  
 KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;  
 KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;  
 KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder;  
 KW PCR primer.

XX  
SO  
Homo sapiens

OS  
XX  
PN

XX  
PD 20-SEP-2001.

28-FEB-2001; 2001WO-US06520.

XX

PR 01-MAR-2000; 2000WO-US05601.

PR 02-MAR-2000; 2000WO-US05841.  
PR 03-MAR-2000; 2000US-197202P

PK	03-MAR-2000; 2000US-187202P.
PR	06-MAR-2000; 2000US-186968P.

PR 14-MAR-2000; 2000US-189320P.

PR 14-MAR-2000; 2000US-189328P.

PR 15-MAR-2000; 2000WO-US06884.  
DP 31-MAR-2000; 2000US-1009385

PR 21-MAR-2000; 2000US-190828P.  
PR 21-MAR-2000; 2000US-191007P

PR 21-MAR-2000; 2000US-191048P;  
21 MAR 2000, 2000US-191007P;

PR 21-MAR-2000; 2000US-191314P.

PR 28-MAR-2000; 2000US-192655P.

PR	29-MAR-2000;	2000US-1930332P.
PR	29-MAR-2000;	2000US-1930333P.
PR	30-MAR-2000;	2000OWO-US08439.
PR	04-APR-2000;	2000US-194449P.
PR	04-APR-2000;	2000US-194547P.
PR	11-APR-2000;	2000US-195375P.
PR	11-APR-2000;	2000US-196000P.
PR	11-APR-2000;	2000US-196187P.
PR	11-APR-2000;	2000US-196690P.
PR	11-APR-2000;	2000US-196820P.
PR	18-APR-2000;	2000US-198121P.
PR	18-APR-2000;	2000US-198585P.
PR	25-APR-2000;	2000US-199397P.
PR	25-APR-2000;	2000US-199550P.
PR	25-APR-2000;	2000US-199654P.
PR	03-MAY-2000;	2000US-201516P.
PR	17-MAY-2000;	2000OWO-US13705.
PR	22-MAY-2000;	2000OWO-US14042.
PR	30-MAY-2000;	2000OWO-US14941.
PR	02-JUN-2000;	2000OWO-US15264.
PR	05-JUN-2000;	2000US-209832P.
PR	28-JUL-2000;	2000OWO-US20710.
PR	22-AUG-2000;	2000US-064448P.
PR	24-AUG-2000;	2000OWO-US23328.
PR	08-NOV-2000;	2000OWO-US30952.
PR	01-DEC-2000;	2000OWO-US32678.
PR	20-DEC-2000;	2000OWO-US34956.

(GETH ) GENENTECH INC.

Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z; Gurney AL;

WPI; 2001-602746/68.

P-PSDB; AAU29104.

—

Novel nucleic acids encoding PRO polypeptides, used to diagnose the presence of tumours, such as prostate and breast tumours, in mammals and to screen for modulators of the compounds -

Claim 2; Fig 161; 774pp: English.

Sequences AAS45925-AAS46231 represent DNA molecules encoding and PCR primers for PRO polypeptides of the invention. The sequences of the invention can be used to detect the presence of a tumour in a mammal by comparing the level of expression of a PRO polypeptide in a test sample of cells from the animal and a control sample of normal cells, whereby a higher level of expression in the test sample indicates the presence of a tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats and rabbits but are preferably human. The polypeptides can be used to stimulate tumour necrosis factor (TNF) alpha release from human blood, when contacted with it. A specific polypeptide can be used to stimulate the proliferation or differentiation of chondrocyte cells. The PRO proteins can be used to determine the presence of tumours and also susceptibility to tumour development, particularly adrenal, lung, colon, breast, prostate, rectal, cervical, or liver tumours, in mammalian subjects. The oligonucleotide probes specific for the PRO nucleic acids can be used for genetic analysis of individuals with genetic disorders.

Sequence 2095 BP; 629 A; 382 C; 418 G; 666 T; 0 other;

Query Match 15.5%: Score 130.4: DB 22: Length 2095.

Best Local Similarity 48.9%; Pred. No. 1.3e-29;

Matches	387;	Conservative	0;	Mismatches	396;	Indels	9;	Gaps	1;
---------	------	--------------	----	------------	------	--------	----	------	----

[illegible]

QY 20 TTACAAGAAAGACGGGAACCTTCCTTAAGCTCCAGATACAGACTGCAGGACACCTC 79

Db 494 TTACAGACAAGACTTTACACTTCACACTTCGAGAGCATTCAAACTGCTCTCATCAAAATC 553

-----

QY 80 CCTTCCTGCTCCTGGTGACCTCATCCACAACAGTTGGCTGAGCGCATGGCCATCC 139

554

DB 554 CATTCTGGTCATTCTGGTGACCTCCCAACCTTCAGATGTGAAGCCAGGCAGGCCATTA 613



Db 974 TTACCAAAACCCATATTTCTACAGGAGTATCTTCAAGGTGTCTCCCTCCATACT 1033  
Qy 551 GTCGCGCACCGCTACGTTTCTTCTGGCGACGTGGCGAGTCAAGTGTACATGCTCTCCA 610  
Db 1034 GCAGTGGGTGGTTATATATGTCAGAGATTTGGTCCCAAGGATCTATGAATGATGG 1093  
Qy 611 AGAGGTCCTCATATTAACCTGGAAGACGTGTTTGGGGCTCTGCTCGAAAGGCTGA 670  
Db 1094 GTCAGGTAACCCATCAAGTTGAAGATGTTTATGTCGGGATCTGTTGAATTTATTAA 1153  
Qy 671 ACATCAGATTGAGGAGCTCCACTCCCGACGCTTTTTCAGGGGGCTTACGCTTCT 730  
Db 1154 AAGTCAATTCATATTCAGAGACACAAATCTTTCTTCTATATAGAATCCATTTGG 1213  
Qy 731 CGTATGCTCTTCAGAGGATCGTGGCGCTGCCACTTCATCAAGCCTCGGACTCTCTGG 790  
Db 1214 ATGCTGTCACTGAGACGTGATTTGCAGCCCATGCTTTCTTCAAGGAGATCATCA 1273  
Qy 791 ACTACTGCGAGG 802  
Db 1274 CTTTGGCAGG 1285  
RESULT 10  
AAF44168  
ID AAF44168 standard; cDNA; 2095 BP.  
XX AC AAF44168;  
XX DT 02-APR-2001 (first entry)  
XX DE Human PRO1074 (UNQ531) nucleotide sequence SEQ ID NO:208.  
XX KW Human; secreted and transmembrane protein; PRO; cytostatic;  
XX KW cell death; cancer; chromosomal mapping; gene mapping; tissue typing;  
XX KW diagnostic assay; ss.  
XX OS Homo sapiens.  
XX PN W0200073454-Al.  
XX PD 07-DEC-2000.  
XX PF 30-MAR-2000; 2000WO-US08439.  
XX PR 02-JUN-1999; 99WO-US12252.  
XX PR 23-JUN-1999; 99US-0141037.  
XX PR 07-JUL-1999; 99US-0143048.  
XX PR 20-JUL-1999; 99US-0144758.  
XX PR 26-JUL-1999; 99US-0145698.  
XX PR 28-JUL-1999; 99US-0146222.  
XX PR 17-AUG-1999; 99US-0149396.  
XX PR 15-SEP-1999; 99WO-US21090.  
XX PR 15-SEP-1999; 99WO-US21547.  
XX PR 08-OCT-1999; 99US-0158663.  
XX PR 30-NOV-1999; 99WO-US28313.  
XX PR 01-DEC-1999; 99WO-US28301.  
XX PR 16-DEC-1999; 99WO-US30095.  
XX PR 20-DEC-1999; 99WO-US30911.  
XX PR 05-JAN-2000; 2000WO-US00219.  
XX PR 03-JAN-2000; 2000WO-US00376.  
XX PR 11-FEB-2000; 2000WO-US03565.  
XX PR 18-FEB-2000; 2000WO-US04341.  
XX PR 22-FEB-2000; 2000WO-US04414.  
XX PR 24-FEB-2000; 2000WO-US04914.  
XX PR 02-MAR-2000; 2000WO-US05004.  
XX PR 15-MAR-2000; 2000WO-US05841.  
XX PR 20-MAR-2000; 2000WO-US07377.  
XX (GETH ) GENENTECH INC.  
XX

PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;  
PI Ferrera N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;  
PI Grimaldi CJ, Gurney AL, Kljavin IJ, Napier MA, Pan J, Paoni NF;  
PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;  
PI Zhang Z;  
XX WPI; 2001-032160/04.  
DR P-PSDB; AAB65208.  
XX PRO polynucleotides used to produce polypeptides used to target  
DR bioactive molecules such as toxins, radiolabels or antibodies, to  
XX specific cells, to cause targeted cell death -  
PS Claim 2; Fig 136; 935pp; English.  
XX The present invention describes human secreted and transmembrane PRO  
CC proteins. The PRO proteins have cytostatic activity. The PRO proteins  
CC can be used for targeted delivery of bioactive molecules, such as  
CC toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide  
CC sequences, and their fragments, can be used as hybridisation probes, in  
CC chromosomal and gene mapping, and in the generation of anti-sense RNA  
CC and DNA. They may also be used to produce transgenic animals which are  
CC used to develop and screen therapeutically useful reagents. The PRO  
CC nucleotide and protein sequence can be used for tissue typing and in  
CC treating cancer. Anti-PRO antibodies can be used in diagnostic assays.  
CC AAF44270 to AAF44470 represent PCR primers and hybridisation probes used  
CC in the isolation of human PRO sequences. AAF44087 to AAF44269 and  
CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein  
CC sequences given in the exemplification of the present invention.  
XX  
XX Sequence 2095 BP; 629 A; 382 C; 418 G; 666 T; 0 other;  
Query Match 15.5%; Score 130.4; DB 22; Length 2095;  
Best Local Similarity 48.9%; Pred. No. 1.3e-29;  
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;  
Qy 20 TTTACAGAGAAGACGGGAACCTTCTTAAGCTCCCGAGATACAGACTCGCAGGCACACCTC 79  
Db 494 TTTACAGAGAAGACTTTTCACITTCACACTTCGAGAGCATCAAACTCTCTCAAAATC 553  
Qy 80 CCTCTCGCTCGTGTGACCTCATCCCAACAGATTTGGCTGAGCGATGCCATCC 139  
Db 554 CATTTCTGTCATCTCTGTCACCTCCACCTTTCAGATGTGAAGCCAGGCACCATTA 613  
Qy 140 GCGACAGCTGGGGAAGAGAGAGGATGTTGAAGGGAAGCAGCTGAAGACATCTTCTCTCC 199  
Db 614 GAGTTACTTGGGTGAAAAAAGTCTTGGTGGGATATGAGGTTCTTACATTTTCTTAT 673  
Qy 200 TGGGGACCCACGAGCTGCAGCGGAACGAAGAGG-----TGGACCGAGGAGCC 250  
Db 674 TAGGCCAAGAGGCTGAAAGGAAGACAAATTTTGGCATTTGCTTAGAGGATGAACACC 733  
Qy 251 AGCGACACGGGGACATTCAGAGGATTTCTTAGAGCTCTATTACATCTGACCCCTGA 310  
Db 734 TCTTTTATGTCATATTCGGACAGATTTTTCAGACATATATTAACCTGACCTTGA 793  
Qy 311 AGACATGATGGGCATAGAATGGGTCCATCGCTTTGTCCTCAGCGGGGCTTTGTGATGA 370  
Db 794 AAACCATATGTCATTCAGGTGGGTAACTGATTTTGGCCCAATGCCAAGTACGTAATGA 853  
Qy 371 AAACAGACTCAGACATGTTTCATCAATGTTGACTACTCTGACTGAAGTCTTCTGAAGAAA 430  
Db 854 AGACAGACACTGATGTTTTCATCAATGTTTGGCAATTTAGTGAAGTATCTTTTAAACCTAA 913  
Qy 431 ACAGAAACACCCAGTCTTTCACCTGGCTTCTTGAACACTCAATGAGTTTCCCATCAGGCAGC 490  
Db 914 ACCACTCAGAGAAGTTTTCACAGGTTATCCTCTAATTTGATAATTTATCTCTATAGAGGAT 973  
Qy 491 CATTCAGCAAGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGATACCCACCATCT 550  
Db 974 TTTACCAAAAAACCCATATTTCTTACAGGAGTATCTTTCAAGGTGTCTCCCTCATACT 1033  
Qy 551 GCTCCGGCACCGGCTACGTTGTTTTCTGGCGACGTGGCGAGTCAAGTGTACATGCTCTCCA 610

Db 1034 GCAGTGGTGGTATATATATGTCACAGATTTGGTGCCACAGGATCTATGAATGATGG 1093  
QY 611 AGAGCGTCCCATACATTAACACTGGAAGAGCGTGTGTTGGGGCTCTGCCCTCGAAGGCTGA 670  
Db 1094 GTCACGTAACCAACCATCAAGTTTGAAGATGTTATGTCGGGATCTGTTTGAATTTATTA 1153  
QY 671 ACATCAGATTGGAGGAGCTCCACTCCAGCGACCTTTTTCAGGGGCGTTACGCTTCT 730  
Db 1154 AAGTGAACATTCATATCCAGAGACACAAATCTTTCTTTCTATATAGAATCCATTTGG 1213  
QY 731 CCGTATGCTCTTCAGGAGGATCGTGGCTGCCACTTCATCAAGCTCGGACTCTCTGG 790  
Db 1214 ATGTCGTCAACTGAGAGCTGATGTCAGCCCATGGCTTTCTTCCAAAGGATCATCA 1273  
QY 791 ACTACTGGCAGG 802  
Db 1274 CTTTGGCAGG 1285  
RESULT 11  
AAH15711  
ID AAH15711 standard; cDNA; 2168 BP.  
XX  
AC AAH15711;  
XX  
XX 26-JUN-2001 (first entry)  
XX Human cDNA sequence SEQ ID NO:14092.  
XX Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.  
XX Homo sapiens.  
XX EP1074617-A2.  
XX  
XX 07-FEB-2001.  
XX  
XX 28-JUL-2000; 2000EP-0116126.  
XX  
XX 29-JUL-1999; 99JP-0248036.  
XX 27-AUG-1999; 99JP-0300253.  
XX 11-JAN-2000; 2000JP-0118776.  
XX 02-MAY-2000; 2000JP-0183767.  
XX 09-JUN-2000; 2000JP-0241899.  
XX  
XX (HELI-) HELIX RES INST.  
XX  
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX WPI; 2001-318749/34.  
XX  
XX Primer sets for synthesizing polynucleotides, particularly the 5602  
XX full-length cDNAs defined in the specification, and for the detection  
XX and/or diagnosis of the abnormality of the proteins encoded by the  
XX full-length cDNAs -  
XX  
XX Claim 8; SEQ ID 14092; 2537pp + CD ROM; English.  
XX  
XX The present invention describes primer sets for synthesizing 5602  
XX full-length cDNAs defined in the specification. Where a primer set  
XX comprises: (a) an oligo-dr primer and an oligonucleotide complementary  
XX to the complementary strand of a polynucleotide which comprises one of  
XX the 5602 nucleotide sequences defined in the specification, where the  
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
XX of an oligonucleotide comprising a sequence complementary to the  
XX complementary strand of a polynucleotide which comprises a 5'-end  
XX sequence and an oligonucleotide comprising a sequence complementary to a  
XX polynucleotide which comprises a 3'-end sequence, where the  
XX oligonucleotide comprises at least 15 nucleotides and the combination of  
XX the 5'-end sequence/3'-end sequence is selected from those defined in  
XX the specification. The primer sets can be used in antisense therapy and

CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to  
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.  
XX  
SQ Sequence 2168 BP; 634 A; 405 C; 452 G; 677 T; 0 other;  
Query Match 15.5%; Score 130.4; DB 22; Length 2168;  
Best Local Similarity 48.9%; Pred. No. 1.3e-29;  
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;  
QY 20 TTTACAAGAAAGACGGGAACCTTCTTAAGTCCACAGATACAGCTGCAGGACACACCTC 79  
Db 571 TTTACAGACAAGACTTTCACTTCAGAGCATTCAAACTGCTCTCATCAAAATC 630  
QY 80 CTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 139  
Db 631 CATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 690  
QY 140 GCGACAGCTGGGGGAAAGAGAGGATGCTGAAGGAAAGACGCTGAAGACATCTCTCTCC 199  
Db 691 GAGTTACTTGGGTGAAAGAAAGCTTGTGGGGATATAGGTTCTTACATTTTCTTAT 750  
QY 200 TGGGACACACACAGCTGCGGCGGAAAGAGG-----TGACACAGGAGGCGC 250  
Db 751 TAGGCCAAGAGCTGAAAGGAAGACAAATGTTGGCATTTCTCTTAGAGGATGAACACC 810  
QY 251 AGCGACACGGGACATTTATCCAGAGGATTTCTTAGACGCTTATTACAAATCTGACCCCTGA 310  
Db 811 TTTCTTATGCTGACATAATCCGACAAAGATTTTGTAGACACATAATAAATACCTGACCTTCA 870  
QY 311 AGACCATGATGGGCATAGATGGTCCATCGCTTTTGTCTCAGCGCGGCTTTGTGATGA 370  
Db 871 AAACCATTTATGGCATTCAGGTGGTAACTGAGTTTTCGCCCAAGCCAAAGTAAATGA 930  
QY 371 AAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGTTCTGAGAAAA 430  
Db 931 AGACAGACACTGATGTTTTCATCAATCTGCAATTTAGTGAAGTATCTTTTAAACCTTAA 990  
QY 431 ACAGAACCAACCAAGTTTTCCTGCTGCTTCTGAAACTCAATGAGTTTCCCATCAGGCAGC 490  
Db 991 ACCACTCAGAGAAGTTTTCACAGGTTATCTCTCTAATTGATAATTATCTCTATAGAGAT 1050  
QY 491 CATTGAGCAAGGTTTGTGCTGAGTAAATCTGAATATCGTGGGACAGGTACCCACATCTCT 550  
Db 1051 TTTACCAAAAAACCCATATTTCTTACCAGGATATCTTTCAAGGTGTCCCTCCATCTACT 1110  
QY 551 GCTCGGCACCGGCTACGTTTCTGCGGACGCTGGGAGTCAAGTGTACAAATGCTCTCCA 610  
Db 1111 GCAGTGGGTTGGTTATATAATGTCAGAGATTTGTGCGCAAGGATCTATGAAATGATGG 1170  
QY 611 AGACGCTCCCATACATTAACTGGAAGACGTTGTTGCGGCTCTGCTCGAAAGGCTGA 670  
Db 1171 GTCAGTAAACCCCATCAAGTTTGAAGATGTTTATGCGGGATCTGTTGAATTTATTA 1230  
QY 671 ACATCAGATTGGAGGAGCTCCACTCCAGCGGACCTTTTTTCCAGGGGCTTACGCTTCT 730  
Db 1231 AAGTGAACATTCATATTCAGAGAGACAAATCTTTTCTTCTATATAGAAATCCTATTGG 1290  
QY 731 CCGTATGCTCTTCCAGGAGGATCGTGGCTGCCACTTCATCAAGCGCTCGGACTCTCTGG 790  
Db 1291 ATGCTGTCAACTGAGAGCGTGTGATTGCGAGCCCATGGCTTTCTTCCAGGAGATCATCA 1350  
QY 791 ACTACTGGCAGG 802  
Db 1351 CTTTGGCAGG 1362



RESULT 12	
AA35710	
ID	AA35710 standard; cDNA; 2189 BP.
XX	AA35710;
XX	27-SEP-1999 (first entry)
XX	Human Dendriac cDNA.
XX	Dendriac; Brainiac-2; human; Notch; immune disorder;
XX	neurological disorder; diagnosis; therapy; ss.
XX	Homo sapiens.
OS	
XX	Key Location/Qualifiers
XX	CDS 435..1394
XX	FT /*tag= a
XX	FT sig_peptide 435..509
XX	FT /*tag= b
XX	FT mat_peptide 510..1391
XX	FT /*tag= c
XX	W09931116-Al.
XX	PN
XX	24-JUN-1999.
XX	PF 17-DEC-1998; 98WO-US27049.
XX	PR 17-NOV-1998; 98US-0108928.
XX	PR 18-DEC-1997; 97US-0068006.
XX	PR 12-MAR-1998; 98US-0077687.
XX	(HUMA-) HUMAN GENOME SCI INC.
XX	PI Ebner R, Endress GA, Florence KA, Rosen CA, Ruben SM;
XX	PI Soppet DR, Yu G;
XX	XX
DR	WPI; 1999-430031/36.
DR	P-PSDB; AAY06461.
XX	Human proteins for treating and detecting immune and neurological disorders
XX	Claim 2; Page 110-112; 132pp; English.
XX	This is the nucleotide sequence of cDNA clone HFVIF40, which codes for human Dendriac (see AAY06461), also called Brainiac-2, a novel member of the Brainiac family. Dendriac cDNA is deposited as ATCC 203056 and ATCC 209627. The cDNA was initially discovered in a cDNA library derived from dendritic cells. Additional clones of the same gene were identified in NTERA2 cell, adult pulmonary tissue, salivary gland, ovary, Caco-2 colon adenocarcinoma, smooth muscle, cerebellum, 8-week-old human mrbyo, haemagiolpericytoma, amygdala, substantia nigra and whole brain cDNA libraries. Dendriac polypeptides are used in the recombinant production of Dendriac polypeptides, especially the mature protein or epitope-bearing fragments. Dendriac and Brainiac proteins are members of the Notch family of proteins involved in the control of cell division. The proteins may be used to detect and/or treat immune system and neurological disorders.
XX	Sequence 2189 BP; 657 A; 402 C; 453 G; 677 T; 0 other;
XX	Query Match 15.5%; Score 130.4; DB 20; Length 2189;
XX	Best Local Similarity 48.9%; Pred. No. 1.4e-29;
XX	Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;
QY	20 TTACAAGAAAGCGGACTTCCTTAAGCTCCCATACAGACTGCAGGACACCTC 79
Db	
Db	571 TTACAGACAGACTTTCACCTTCGAGAGCAATTCAAACTGCTCTCATCAAAATC 630
QY	80 CCTTCCTCGTCTGCTGAGCTCATCCACAAACAGTTGGTGCAGCGCATGCCATCC 139

Db	631 CATTTCTGCTCATTTCTGTTGACCTCCACCCCTTCAGATGTAAAGCCAGGAGCCCATTA 690
QY	140 GGCAGACGTGGGGAAGAGAGATGGTGAAGGGAAGAGAGCTGAACACATCTTCTCTCC 199
Db	691 GAGTTACTTGGGGTGAAGAAAGTCTTGGTGGGATATGAGGTTCTTACATTTTCTTAT 750
QY	200 TGGGACACCACGACAGTGCAGCGGAAACGAAAGAGG-----TGGACAGGAGGCC 250
Db	751 TAGCCCAAGAGGCTGAAAGGAAGACAAATGTTGGCATGTCTCTAGAGATGAACACC 810
QY	251 AGCAGACGGGGACATTATCCAGAGGATTTCCCTAGACGCTTATTACAATCTGACCTGA 310
Db	811 TTTCTTTATGGTGACATAATCCGACAAGATTTTTCAGACACATATAATAACCTGACCTGA 870
QY	311 AGACCATGATGGGCATAGAATGGGTCCTCATCGCTTTTGTCTCAGGCGGGCTTTGTGATGA 370
Db	871 AAACCATATATGGCATTCAGGTGGTAACTGAGTTTTCGCCCAATGCCAAGTACGTAATGA 930
QY	371 AAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACCTCTCTCTGAAGAAA 430
Db	931 AGACAGACACTGATGTTTTCATCAATACTGGAATTTAGTGAAGTATCTTTTAAACCTAA 990
QY	431 ACAGAACAAACGAGTTTTCCTACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGAGC 490
Db	991 ACCACTCAGAGAAGTTTTCACAGGTTATCTCTTAATGATTAATTTCTCTATAGAGAT 1050
QY	491 CATTCAGCAAGTGGTTTGTCTCAGTAATCTGAATATCCGTGGGACAGGTACCCACCATCT 550
Db	1051 TTTACCAAAAACCCATATTTCTTACCAGAGTATCTCTTCAAGGTGTTCCCTCCCATCT 1110
QY	551 GCTCCGGCACCGGCTACGTGTTTCTGCGACGTGCGAGTCAGGTGTACAAATGCTCTCCA 610
Db	1111 GCAGTGGTGGGTATATAATGTCAGAGATTTGGTCCAAGGATCTATGAATGATGG 1170
QY	611 AGAGCGTCCCATACATTAACCTGGAAGACGTGTTTGTGGGGCTCTGCTCGTGAAGGCTGA 670
Db	1171 GTCACGTAAACCCCATCAAGTTTGAAGATGTTTATGTCGGGATCTGTTGAATTTATTA 1230
QY	671 ACATCAGATGGAGAGCTCCACTCCAGCCGACACCTTTTCCAGGGGGCTTACGCTTCT 730
Db	1231 AAGTGAACATTCATATTCAGAGAAGACAAATCTTTTCTTCTATATAGAACATCCATTGG 1290
QY	731 CCGTATGCTCTTCAGGAGGATCGTGGCTGCCACTTCATCAAGCTCGGACTCTCTGG 790
Db	1291 ATGCTCTCAACTGAGAGCTGTGATTCAGGCCCATGGCTTTTCTTCCAGGAGATCATCA 1350
QY	791 ACTACTGGCAGG 802
Db	1351 CTTTTTGGCAGG 1362
XX	RESULT 13
XX	AA35710
ID	AA35710 standard; cDNA; 1266 BP.
XX	AA35710;
XX	09-JUL-1999 (first entry)
XX	cDNA encoding a protein identified by the signal sequence trap method.
XX	Signal sequence trap method; SST method; immunisation; inhibition;
XX	infection; allergy; cancer; regulation; tissue formation; tissue repair;
XX	activin activity; inhibin activity; chemokine activity; agonist;
XX	cytokine activity; blood coagulation regulation; immune disorder;
XX	metabolic disorder; hormonal disorder; immune disorder;
XX	severe combined immunodeficiency; SCID; AIDS; thrombosis; cancer;
XX	wound; ss.
OS	Homo sapiens.
XX	W09918126-Al.





Db 826 GCGNGGNGGNGGNTAYGTNATGWSNMGNCNACNGTNNMGNGNYTNCARGCNATHATG 885  
QY 610 AAGAGCGTCCCATACATTAAACTGGAAGACGCTGTTTGTGGGGCTCTGCGCTCGAAGGCTG 669  
Db 886 GARGAYCNGARYTNYTNWSNATHGAYGAYGNTTYGTNGGNATGTGYTNNMGNGNYTN 945  
QY 670 AACATCAG 677  
Db 946 GGNYTNS 953

Search completed: April 11, 2003, 20:07:43  
Job time : 159.586 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 15:35:54 ; Search time 935.163 Seconds  
(without alignments)  
14547.424 Million cell updates/sec

Title: US-09-914-152-2\_COPY\_492\_1331  
Perfect score: 840  
Sequence: 1 ttcaagaacagctcttctgt.....gggaagattgtccgcctgtc 840

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- EST:\*
- 1: em\_estba:\*
  - 2: em\_esthum:\*
  - 3: em\_estin:\*
  - 4: em\_estmu:\*
  - 5: em\_estov:\*
  - 6: em\_estopl:\*
  - 7: em\_estro:\*
  - 8: em\_htc:\*
  - 9: gb\_est1:\*
  - 10: gb\_est2:\*
  - 11: gb\_htc:\*
  - 12: gb\_est3:\*
  - 13: gb\_est4:\*
  - 14: gb\_est5:\*
  - 15: em\_estfun:\*
  - 16: em\_estom:\*
  - 17: gb\_gss:\*
  - 18: em\_gss\_hum:\*
  - 19: em\_gss\_inv:\*
  - 20: em\_gss\_pln:\*
  - 21: em\_gss\_vrt:\*
  - 22: em\_gss\_fun:\*
  - 23: em\_gss\_mam:\*
  - 24: em\_gss\_mus:\*
  - 25: em\_gss\_other:\*
  - 26: em\_gss\_pro:\*
  - 27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	366.6	43.6	434	12 BF748181	BF748181 RC3-BN041
c 2	252.6	30.1	338	9 AJ003597	AJ003597
c 3	233	27.7	400	17 AZ768940	AZ768940 1M0569K09
c 4	184	21.9	1101	17 CNS05BT1	AL330238 Tetraodon
5	144.8	17.2	636	13 BG965086	BG965086 602829157
6	135	16.1	1514	11 AK003837	AK003837 Mus muscu

7	132.2	15.7	596	13	BJ501651
8	128.8	15.3	3564	11	BC028571
9	127	15.1	233	10	BB595953
10	123.4	14.7	444	13	BM403751
11	122.6	14.6	879	14	BQ895238
12	117.2	14.0	965	17	CNS051B2
c 13	115.4	13.7	1021	17	CNS044T4
c 14	114.8	13.7	574	10	BF652985
15	109.4	13.0	1101	13	BM548138
16	108	12.9	620	9	AUI33606
17	107.6	12.8	712	9	AL042887
c 18	101.4	12.1	966	17	CNS04N8R
19	101	12.0	921	17	CNS03CTW
20	99.8	11.9	605	14	BQ417019
21	99.8	11.9	815	10	AW128625
c 22	97.2	11.6	464	9	AA972783
c 23	95.6	11.4	611	13	BJ039496
c 24	95.4	11.4	411	9	AA858037
c 25	93.8	11.2	793	12	BG207694
26	93.4	11.1	675	10	BB648067
c 27	92.6	11.0	809	17	CNS04BGI
c 28	90.8	10.8	625	13	BJ054223
c 29	90.8	10.8	704	13	BM315483
c 30	90.6	10.8	583	13	BI411305
31	87.6	10.4	1963	11	AK008674
32	83.6	10.0	923	9	AU067264
33	83.4	9.9	983	10	BB609433
34	82.8	9.9	584	10	BE373780
35	80.2	9.5	360	9	AUI17194
36	76.2	9.1	1041	17	CNS03BMU
37	70.6	8.4	427	13	BI476269
38	69.6	8.3	581	17	AZ381343
c 39	68	8.1	1101	17	CNS05G62
c 40	68	8.1	1568	11	AK007600
41	66.4	7.9	675	12	BG077355
42	65.8	7.8	647	9	AL775073
c 43	65.6	7.8	909	9	AL522683
c 44	64.8	7.7	933	14	BQ926448
45	64.8	7.7	950	14	BQ946920

ALIGNMENTS

RESULT 1  
BF748181/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

BF748181  
RC3-BN0411-021000-021-c04 BN0411 Homo sapiens cDNA, mRNA sequence.  
BF748181  
BF748181.1 GI:12074857  
EST.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Naga,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.

434 bp  
linear  
EST 10-JAN-2001

TITLE  
JOURNAL  
MEDLINE  
COMMENT

Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001

```

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?l1=RC3&l2=RC3-BN0411-021000-021-c04&t3=2000-10-02&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 24
High quality sequence stop: 432.
Location/Qualifiers
    1. .434
       /organism="Homo sapiens"
       /db_xref="taxon:9606"
       /clone_lib="BN0411"
       /dev_stage="Adult"
       /note="Organ: breast_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
104 a      117 C      130 G      83 T
                                     BASE COUNT
                                     C      G      A      T

```

Query Match	43.6%	Score 366.6	DB 12	Length 434
Best Local Similarity	98.9%	Pred. No. 5.5e-94		
Matches 369	Conservative 0	Mismatches 4	Indels 0	Gaps 0
QY 468	CAATGAGTTCCTCCCATCAGGCAGCCATTTCAGTAAATCTGAATATCC	527		
Db				
QY 434	CAATGAGTTCCTCCCATCAGGCAGCCATTTCAGTAAATCTGAATATCC	375		
Db				
QY 528	GTGGGACAGGTACCCACCATTCGCTCGGCACCGGCTACGTGTTTCTGGCGACGTGGC	587		
Db				
QY 374	GTGGGACAGGTACCCACCATTCGCTCGGCACCGGCTACGTGTTTCTGGCGACGTGGC	315		
Db				
QY 588	GAGTCAGGTGTACAATGTCTCCAGAGCGTCCCATACATTAATCGAAGACGTGTTGT	647		
Db				
QY 314	GAGTCAGGTGTACATGTCTCCAGAGCGTCCCATACATTAATCGAAGACGTGTTGT	255		
Db				
QY 648	GGGGCTCTGCTCGAAAGGCTGAACATCAGATTGGAGGAGCTCCACATCCACGCGACCTT	707		
Db				
QY 254	GGGGCTCTGCTCGAAAGGCTGRACATCAGATTGGAGGAGCTCCACATCCACGCGACCTT	195		
Db				
QY 708	TTTTCCAGGGGGCTTAGCGTTCTCCGTTATGCCCTTTCAGGAGGATCGTGGCTGCCACTT	767		
Db				
QY 194	TTTTCCGCGGGGGCTTAGCGTTCTCCGTTATGCCCTTTCAGGAGGATCGTGGCTGCCACTT	135		
Db				
QY 768	CATCAGGCTTCGACTCTCTTGACTACTGCGAGGCTCTACAGAAATCCCGGGGGGGA	827		
Db				
QY 134	CATCAGGCTTCGACTCTCTTGACTACTGCGAGGCTCTACAGAAATCCCGGGGGGGA	75		
Db				
QY 828	TTGTCGCGCTGTC	840		
Db				
QY 74	TTGTCGCGCTGTC	62		
Db				

RESULT 2	
AJ003597/c	
LOCUS	
DEFINITION	338 bp mRNA linear EST 04-DEC-1997 AJ003597 Selected chromosome 21 cDNA library Homo sapiens cdNA clone WPIp19-12J9, mRNA sequence.
ACCESSION	AJ003597
VERSION	AJ003597.1 GI:2578270
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 338)
AUTHORS	Szulzewsky,I., Hunt,E., Nguyen,M., Korn,B., Roehrdanz,B., Lehrach, H., and Yaspo,M.L.H.
TITLE	An integrated transcript map for the whole human chromosome 21

```

JOURNAL      Unpublished (1997)
COMMENT      Contact: Vaspö, M.-L.
              Max Planck Institut fuer Molekulare Genetik
              Innestrasse 73, D14195 Berlin-Dahlem, Germany.
FEATURES     Location/Qualifiers
              1..338
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /map="21q"
              /clone="MP19-12J9"
              /clone_lib="Selected chromosome 21 cDNA library"
              /note="Swaroop et al. (1991) Nucleic Acids Res. 19: 1954."
BASE COUNT   76 a      92 c      82 g      87 t      1 others
ORIGIN
Query Match  30.1%; Score 252.6; DB 9; Length 338;
Best Local Similarity 98.5%; Pred. No. 2.3e-61;
Matches 255; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy  1  TTCAAGACAGTCCTTTGTTTACAAAGAACGGGAACTCCTTAAGCTCCAGATACA 60
    |||
Db  259 TTCAAGAACAGTCCTTTGTTTACAAAGAACGTGAACCTCCTTAAGCTCCAGATACA 200
    |||

Qy  61  GACTGCAGGCAGACACTCCCTTCCTCGTCTGCTGCTGAGCTCATCCACAAACAGTTG 120
    |||

Db  199 GACTGCAGGCAGACACTCCCTTCCTCGTCTGCTGAGCTCATCCACAAACAGTTG 140
    |||

Qy  121 GCTGAGCGCATGCCATCCGGCAGACGTGGGGAAAGAGAGGATGCTGAAGGGAAAGCAG 180
    |||

Db  139 GCTGAGCGCATGCCATCCGGCAGACGTGGGGAAAGAGAGGAGCGGTGAAGGGAAGCGAG 80
    |||

Qy  181 CTGAGACATCTTCTCTCTGGGACACACAGCAGTGCAGCGGGAACGAAAGAGGTGGAC 240
    |||

Db  79  CTGAAGACATCTTCTCTCTGGGACACACAGCAGTGCAGCGGGAACGAAAGAGGTGGAC 20
    |||

Qy  241 CAGGAGAGCCAGCGACAG 259
    |||

Db  19  CAGGAGAGCCAGCGACAG 1
    |||

RESULT 3
A2768940/c
LOCUS      A2768940
DEFINITION 1M0569K09F Mouse 10kb plasmid UUGCIM library Mus musculus genomic
clone UUGCIM0569K09 F, DNA sequence.
ACCESSION A2768940
VERSION    A2768940.1 GI:12888559

```

RESULT 3  
 AZ768940/c  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

AZ768940  
 400 bp DNA linear GSS 16-FEB-2001  
 1M0569K09F Mouse 10kb plasmid UUGCIM library Mus musculus genomic  
 clone UUGCIM0569K09 F, DNA sequence.  
 AZ768940  
 GSS.  
 AZ768940.1 GI:12888559  
 house mouse.  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 400)  
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
 ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.,  
 and Wright,D.,Weiss,R.  
 Mouse whole genome scaffolding with paired end reads from 10kb  
 plasmid inserts  
 Unpublished (2000)  
 Contact: Robert B. Weiss  
 University of Utah Genome Center  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT,  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0569 row: K column: 09  
 Seq primer: CGTTGTAAACGACGCCCACT  
 Class: plasmid ends  
 High quality sequence stop: 400.

ORGANISM	SEQUENCE	POSITION	SCORE	IDENTITY	LENGTH	MISSING	INDELS	GAPS	OTHERS
Tetraodon nigroviridis.									
Tetraodon nigroviridis									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;									
Acanthomorphia; Acanthopterygii; Percomorpha; Tetraodontiformes;									
Tetraodontidae; Tetraodon.									
1 (bases 1 to 1101)									
Roest-Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,									
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,									
Saurin,W. and Weissenbach,J.									
Human gene number estimate provided by genome wide analysis using									
Tetraodon nigroviridis DNA sequence									
Unpublished									
2 (bases 1 to 1101)									
Roest-Crolius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,									
Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and									
Weissenbach,J.									
Characterization and repeat analysis of the compact genome of the									
freshwater pufferfish Tetraodon nigroviridis									
Unpublished									
3 (bases 1 to 1101)									
Genoscope.									
Direct Submission									
Submitted (12-APR-2000)									
This sequence is a single read and was generated as part of a large									
scale clone-end sequencing project of the Tetraodon nigroviridis									
genome. For more information, please take a look at									
http://www.genoscope.cns.fr/Tetraodon.									
Location/Qualifiers									
1..1101									
/organism="Tetraodon nigroviridis"									
/db_xref="taxon:9983"									
/clone="013P23"									
/clone_lib="C"									
/note="Genoscope sequence ID : C0CC013CH121-end : T3"									
BASE COUNT	235 a	239 c	325 g	283 t	19 others				
ORIGIN									
Query Match	21.9%;	Score 184;	DB 17;	Length 1101;					
Best Local Similarity	58.5%;	Pred. No. 1.9e-41;							
Matches 354;	Conservative 0;	Mismatches 242;	Indels 9;	Gaps 2;					
QY	77	CTCCCTTCCTCGTCCTGGTGGTACCTCCACAAACAGTTGGCTGAGCGCATGGCCA	136						
Db	663	CTCCTTTCCTGGTATCTCTCATGACACCGACAGGAGTTTACGCGCGCAGGCCA	604						
QY	137	TCCGGCAGCTGGGGGAAGAGGAGTGGTGAAGGGAAAGCAGCTGAAGACATTCTTC	196						
Db	603	TCCGGAGACCTGGGGGAGCAGACACCTTCACCGAGCTCCGATTCACCGCTTTC	544						
QY	197	TCCTGGGACACACAGCAGTGCAGCGGAAACGAAAGAGTGGACCGAGAGCCAGCGAC	256						
Db	543	TNCTGGCAGGAACACCGATGAAGTCTCAACAGATGGTGGAGCAGGAGCCATCT	484						
QY	257	ACGGGACATTCAGAGAGATTTCCTAGACGTCATTATTAATCTGACCCCTGAAGACCA	316						
Db	483	TCCATGATCTGTATGGAGAACTTCATCGATTCTACCACAACTCAGCTCAAAACTT	424						
QY	317	TGATGGGATACAATGGGTCCATCGCTTTTGTCTCAGCGCGCGTTGTGTATGAACACAG	376						
Db	423	TGATGGGATCGCTGGGTGGCCACCTTCGCCCAAAGCTCAATATGTATGATGAAGACAG	364						
QY	377	ACTCAGACATGTTCAATCAATGTTGACTATCTGACTGAACCTCTCTCTGAAG-----	430						
Db	363	ACAGGACATCTCGTCAACATCNACAATCTNATCTCAAGCTCTGAAGCCAGCACCA	304						







AUTHORS Kohara,Y., Shin-I.T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.  
TITLE Medaka EST Project in Takeda's lab  
JOURNAL Unpublished (2001)  
COMMENT Contact: Tadasu Shin-I  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.

FEATURES

source  
1...596  
/organism="Oryzias latipes"  
/strain="d-IR"  
/db\_xref="taxon:8090"  
/clone="MF01FSA049L05"  
/clone\_lib="MF01FSA\_CONA"  
/sex="mixture of female and male"  
/tissue\_type="whole embryo"  
/dev\_stage="fry stage 40"  
BASE COUNT 138 a 168 c 151 g 137 t 2 others  
ORIGIN

Query Match 15.7%; Score 132.2; DB 13; Length 596;  
Best Local Similarity 55.2%; Pred. No. 9.8e-27;  
Matches 277; Conservative 0; Mismatches 222; Indels 3; Gaps 1;

Qy 177 GCAGCTGAAGACATCTCTCTCTGGGACACAGAGTGCACGGGAAACGAAGAGT 236  
Db 13 GCGAGTGTACCCCTTTCTTGTGGGGCGCAGCATGGCGCTCTCTGAACACAGTGT 72  
Qy 237 GGACGAGGAGCGCCGACGACGGGACATTTCCAGGAAGATTTCCTAGACGCTATTA 296  
Db 73 GGACGAGGAGTGCAGATCTTCATGATGTAGTCTGGAGGACTTATGTACTGTACCA 132  
Qy 297 CAATCTGACCCCTGAAGACATGATGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGGC 356  
Db 133 CAACCTGACGCTNAAGACCCCTGATGGGTATCGCGTGGGTGGCTACTCTACCAAGGC 192  
Qy 357 GCGTTTGTGATGAACACAGACTCACACATGTTTCATCAATGTTGACTATCTGACTGAAC 416  
Db 193 CCAATATGATCTCAAGACGAGCAGTGCACATCTTTGTCAACATGGAGAACCTCATCTACAA 252  
Qy 417 GCTTCTGAAG---AAAAAGACAACACCCAGGTTTTCACCTGGGTTCTTGAACATCAATGA 473  
Db 253 CTTCTGAAGCCCAACCAACGCTAGGAGAAGTATTTTACAGGTTACGTCATCAACGG 312  
Qy 474 GTTCCCATCAGGCGCCATTTCAGCAAGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGA 533  
Db 313 GGGCCGATCAGACATACGCGCAAGTGGTACATGCCCGGGATCTTTATCCGACAG 372  
Qy 534 CAGGTACCCACCATCTGTCTCGGCAACGGCTACGTTGTTTCTGGCGACGTTGGCGAGTCA 593  
Db 373 CAAATACCCACCTCTGTCTCGGCAACGGCTACGTTCTCTCAGCAGATGTAGCTGAGCT 432  
Qy 594 GGTGTACATGCTCCAGAGGCTCCCATACATTAACTGGAGACGCTGTTTGTGGGGCT 653  
Db 433 CATATACNAGATCTCTTTGCACACACAGGCTCTCCACCTGGAGGACGCTGATGTTGGAGT 492  
Qy 654 CTGCTCGAAGGCTGAACATC 675  
Db 493 TTGCTCGTAGCTGGGCATC 514

RESULT 8  
LOCUS BC028571 3564 bp mRNA linear HTC 21-AUG-2002  
DEFINITION Homo sapiens, UDP-Gal:betaGalNAc beta 1,3-galactosyltransferase, polypeptide 3, clone IMAGE:4838965, mRNA.  
ACCESSION BC028571  
VERSION BC028571.1 GI:22382221  
KEYWORDS HTC.  
SOURCE human.

ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 3564)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Strausberg,R.  
Direct Submission  
Submitted (23-APR-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC project URL: http://mgc.nci.nih.gov  
Contact: MGC help desk  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Institute for Systems Biology  
http://www.systemsbio.org  
contact: amadan@systemsbiology.org  
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha  
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

REMARK  
COMMENT

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Series: IRAC Plate: 34 Row: P Column: 23  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 15451873  
This clone has the following problem: incomplete processing.

FEATURES

source  
1...3564  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4838965"  
/tissue\_type="Testis"  
/clone\_lib="NIH\_MGC\_97"  
/lab\_host="DH10B"  
/note="Vector: pBluescript"  
BASE COUNT 1048 a 661 c 741 g 1114 t  
ORIGIN

Query Match 15.3%; Score 128.8; DB 11; Length 3564;  
Best Local Similarity 48.7%; Pred. No. 2.2e-25;  
Matches 386; Conservative 0; Mismatches 397; Indels 9; Gaps 1;

Qy 20 TTTACAGAAGACGCGGAACCTTCCTTAAGCTCCAGATACACACTGCAGGCACACCTC 79  
Db 864 TTTACAGAAGACTTTTCATCTCAGACTTCGAGAGCATTCACAACTGCTCTCAATAATC 923  
Qy 80 CCTTCTCGCTCGTGGTGACCTCATCCACAAACAGTTGGCTGAGCGCATGCCATCC 139  
Db 924 TATTTCTGGTCACTTCTGTGACCTCCACCCCTTCAGATGTGAAGCCAGGCGCATTA 983  
Qy 140 GGCAGACGTGGGGAAGAGAGGATGGTGAAGGGAAGCAGCTGAAGACATTTCTCTCC 199  
Db 984 GAGTTACTTGGGTGAAAAAAGTCTTGGTGGGATATGAGTTCTTACATTTTCTTAT 1043  
Qy 200 TGGGGACCCACGACGTCGACGGGAAACGAAGAGG-----TGGACCGAGGAGCC 250  
Db 1044 TAGGCCAAGAGGCTGAAAAGGAGACAAAATGTTGGCATTTGCTTAGAGGATGAACCC 1103  
Qy 251 AGCGACACGGGGACATTATCCAGAAGGATTTCTTAGAGCTCTATTACATCTGACCCCTGA 310  
Db 1104 TTCTTTATGTGACATAATCCGACAGATTTTTTGGACACATATATAACCTGACCTTGA 1163  
Qy 311 AGACCATGATGGGCATAGAATGGGTCCATCGCTTTTGTCTCAGCGGGGTTGTGATGA 370  
Db 1164 AAACCATTTAGCATTCAGTGGGTACTGAGTTTGGCCCAATGCCAAGTACGTAATGA 1223  
Qy 371 AAACAGACTCAGACATGTTTCATCAATGTTGACTACTCTGACTCAACTGCTTCTGGAAGAAA 430  
Db 1224 AGACAGACACTGATGTTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAA 1283

QY 431 ACAGAACACACAGGTTTTCACCTGGCTCTTGAACACTCAATGAGTTTCCCATCAGGCAGC 490  
 Db 1284 ACCACTCAGAGAAGTTTTCACAGGTTATCCTCTAATGATTAATATTCCTATAGAGGAT 1343  
 QY 491 CATTCAGCAAGTGGTTGTTCAGTAAATCTGAATATCGTGGACAGGTACCCACCAATCT 550  
 Db 1344 TTACCAAAAACCCATATTTCTACAGAGATATCCTTCAAGGTGTTCCTCCATACT 1403  
 QY 551 GCTCGGACACCGCTACGCTTTTCTGGCAGCTGGGAGTCAAGTGTACAAATGTCTCCA 610  
 Db 1404 GCAGTGGGTTGGTTATATAATGTCAGAGATTTGGTCCCAAGATCTATGAATGATGG 1463  
 QY 611 AGAGCGTCCATACATTAACTGGAAGACGTGTTTGGGGCTCTGCTCGAAAGGCTGA 670  
 Db 1464 GTCAGCTAAACCCATCAAGTTTGAAGATGTTTATGTGCGGATCTGTTGAATTATTA 1523  
 QY 671 ACATCAGATTGAGAGCTCCACTCCGCGACCTTTTTCAGGGGCTTACGCTTCT 730  
 Db 1524 AAGTGAATTCATATTCAGAGACACAAATCTTTTCTTATATATAGATCCATCTGG 1583  
 QY 731 CGTATGCTCTTTCAGGAGGATCGTGGCTGCCACTTCATCAAGCCTCGGACTCTCTGG 790  
 Db 1584 ATGCTGTCAACTGAGACGTGTGATTCAGCCCATGCTTTCTTCCAGGAGATCATCA 1643  
 QY 791 ACTACTGGCAGG 802  
 Db 1644 CTTTTTGGCAGG 1655

RESULT 9  
 BB595953 233 bp mRNA linear EST 30-NOV-2000  
 LOCUS BB595953 RIKEN full-length enriched, 0 day neonate cerebellum Mus  
 DEFINITION musculus cDNA clone C230001K09 5', mRNA sequence.

ACCESSION BB595953  
 VERSION BB595953.1 GI:11492555  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus

REFERENCE  
 AUTHORS Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 1 (bases 1 to 233)  
 Aizawa, K., Akahira, S., Akimura, T., Arai, A., Arakawa, T., Carninci, P., Hanagaki, T., Hayatsu, N., Hiraoka, T., Hirozane, T., Hodojima, Y., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kawai, J., Kojima, Y., Konno, H., Kusakabe, M., Matsuyama, T., Miyazaki, A., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Okazaki, Y., Okido, T., Owa, C., Sakai, C., Sakai, K., Sakai, D., Sato, K., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Tanaka, T., Toya, T., Watahiki, A., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshiki, A., Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Mouse ESTs (Aizawa, K. et al. 2000)  
 Unpublished (2000)

TITLE  
 JOURNAL  
 COMMENT

Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
 The Institute of Physiological and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gs.riken.go.jp/  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sakai, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Thermotabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
 Itoh, M., Kitzunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.  
 Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)  
 Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)  
 please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

#### FEATURES

source Location/Qualifiers  
 1..233  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="C230001K09"  
 /clone\_lib="RIKEN full-length enriched, 0 day neonate cerebellum"  
 /tissue\_type="cerebellum"  
 /dev\_stage="0 day neonate"  
 /lab\_host="DH10B"  
 /note="Site\_1: Sali; Site\_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'-GAGAGAGAGATCCAGAGCTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 479.0. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGATTCGAGTTAATAATTAATCCGCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 57 a 57 c 74 g 44 t 1 others  
 ORIGIN

Query Match 15.1%; Score 127; DB 10; Length 233;  
 Best Local Similarity 71.6%; Pred. No. 1.9e-25;  
 Matches 166; Conservative 0; Mismatches 66; Indels 0; Gaps 0;  
 QY 79 CCCTTCCTCGTCTGCTGACCTCATCCACAAACAGTTGGCTGAGCGCATGGCCATC 138  
 Db 2 CCTTTCCTGCTGCTGCTGACCTCATCCAAAGCAGCTGGCGCTCGCATGGCCATC 61  
 QY 139 CGCAGACGCTGGGGAAAGAGAGATGGTGAAGGAAAGCAGCTGAAGACATCTTCCTC 198  
 Db 62 CGGAAGAAGTGGGTAAGAGACATCTGTGGAGGGCCACAGGTGAGGACCTCTTCCT 121  
 QY 199 CTGGGACACACAGCAGCTGCGGGGAAACGAAAGAGTGGACACAGGAGGAGGACAC 258  
 Db 122 CTGGGATCTCCGACAGGACCGAGAAATGGACGCCACACCGTGGAAAGCAGNAGTAC 181  
 QY 259 GGGGACATTATCCAGAGGATTTCTAGAGCTTATTACAATCTGACCCCTGA 310  
 Db 182 GGGATATTATCCAGAGGATTTCAAGGATGCTTCTCAACGTCGACCCCTGA 233

RESULT 10  
 BB403751 444 bp mRNA linear EST 22-JAN-2002  
 LOCUS BB403751 zebrafish skeletal muscles cDNA library Danio rerio cDNA  
 DEFINITION clone zam6155 5', mRNA sequence.  
 ACCESSION BB403751  
 VERSION BB403751.1 GI:18255132  
 KEYWORDS EST.  
 SOURCE zebrafish.  
 ORGANISM Danio rerio  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.  
 1 (bases 1 to 444)  
 Ton, C., Dempsey, A., Hwang, D.M. and Liew, C.C.  
 Identification and Characterization of Expressed Sequence Tags from Zebrafish Skeletal Muscles cDNA Library  
 Unpublished (2002)



```
LOCUS      CNS051B2      965 bp      DNA      linear      GSS 26-MAY-2000
DEFINITION Tetraodon nigroviridis genome survey sequence T3 end of clone
            045P20 of library A from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION  AL338663.1 GI:8232421
VERSION    AL338663.1
KEYWORDS   GSS: genome survey sequence.
SOURCE     Tetraodon nigroviridis.
ORGANISM   Tetraodon nigroviridis.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE  1 (bases 1 to 965)
AUTHORS   Bernot,A., Fizames,C., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Saurin,W. and Weissenbach,J.
TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 965)
AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
TITLE     Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
JOURNAL   Unpublished
REFERENCE  3 (bases 1 to 965)
AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
TITLE     Direct Submission
JOURNAL   Submitted (12-APR-2000)
COMMENT   This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES   Location/Qualifiers
            source
            1..965
            /organism="Tetraodon nigroviridis"
            /db_xref="taxon:99883"
            /clone="045P20"
            /clone_lib="A"
            /note="Genoscope sequence ID : C0AA045DH10A1-end : T3"
BASE COUNT 249 a 245 c 220 g 229 t 22 others
ORIGIN
Query Match 14.0%; Score 117.2; DB 17; Length 965;
Best Local Similarity 54.6%; Pred. No. 2.5e-22;
Matches 242; Conservative 0; Mismatches 199; Indels 2; Gaps 1;
QY 234 GTGGACGAGGAGCGGACGACGCGGACATTTATCCAGAAGGATTTCTAGACGCTA 293
DB 42 GCTGGAGGAGAGATCAGATCTTTCACGACATCGTTGTGGAGGATTTTNGACTCGTA 101
QY 294 TTACAATCTGACCCCTGAAGACCATATGGGCATAGAAATGGTCCATCGCTTTTGTCTCA 353
DB 102 CCACAATTGACATCAAGACCCCTGATGGCATGCGCTGGTGGTACGTTTGTCTCAA 161
QY 354 GCGGGCGTTTGATGAAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGA 413
DB 162 GGCTCAGTATGCTNCTAAAGACGAGAGTGCATCTTTGTCAACATGGAACCTCATCTT 221
QY 414 ACTGCTTCTGAAGAAAACAGAACACACAG--GTTTTTCACTGGCTTCTTGAACACTCAAT 471
DB 222 TAACCTCTGAAACCCAAACACCAACCCAGGAGGAGGTACTTCACTGGTTATGNATCAA 281
QY 472 GAGTTTCCATCAGCGCCATTCAGCAAGTGGTTTGTAGTAAATCTGAATATCCGTGG 531
DB 282 TGGTGGCCCAATNAGAGACATCGGCANAAAGTGTAGATGTCNAGAGATCTGTACCCAGAG 341
QY 532 GACAGGTACCCACCATTTCTGTCGCGACCGGCTAGCTGTTTCTGGGACGCTGGCGAGT 591
DB 342 AGCAAGINCCGCGCCCTTTGTTCCGGCACTGTTAGCTTCTTCCGCGGACGCTGGCGGAG 401
```

```
QY 592 CAGGTGTAAATGTCTCCAAGACGCTCCCATACATTAAACTGGAAGACGTTTGTGGG 651
DB 402 CTTATTTTTTATNCCTCTTACACACAGGCTCTGCACCTCGAGGACGTTATNGGT 461
QY 652 CTCCTGCCTCGAAGCGCTGAACAT 674
DB 462 NTGTGCCTTCGCAAGCTGGCAT 484
RESULT 13
CNS044T4/c
LOCUS
DEFINITION Tetraodon nigroviridis genome survey sequence GSS 18-MAY-2000
            082009 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION  AL274513.1 GI:7995991
VERSION    AL274513.1
KEYWORDS   GSS: genome survey sequence.
SOURCE     Tetraodon nigroviridis.
ORGANISM   Tetraodon nigroviridis.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE  1 (bases 1 to 1021)
AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 1021)
AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
TITLE     Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
JOURNAL   Unpublished
REFERENCE  3 (bases 1 to 1021)
AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
TITLE     Direct Submission
JOURNAL   Submitted (12-APR-2000)
COMMENT   This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES   Location/Qualifiers
            source
            1..1021
            /organism="Tetraodon nigroviridis"
            /db_xref="taxon:99883"
            /clone="082009"
            /clone_lib="G"
            /note="Genoscope sequence ID : C0BG082AH05SP1-end :
            PUC-Ori"
BASE COUNT 207 a 262 c 290 g 259 t 3 others
ORIGIN
Query Match 13.7%; Score 115.4; DB 17; Length 1021;
Best Local Similarity 59.8%; Pred. No. 8.3e-22;
Matches 214; Conservative 1; Mismatches 138; Indels 6; Gaps 1;
QY 62 ACTGCAGGACACACCTTCCTCGCTGCTGGTGCCTCATCTCCCTCATCCCCACAACAGTTGG 121
DB 365 AGTGCAGGAGAGCTCCCTTCCTCATCTCCCTCATCGTGCAGAACCCAGCCGCGTGG 306
QY 122 CTGAGCGCATGGCCATCCGGCAGACGTGGGGAAGAGAGGATGTTGAAGGAAAGCAGC 181
DB 305 ATGCCGCTAACGCCATCCGCCACACGTGGGAAACGAGACACACGATGGCGCTGGGCT 246
QY 182 TGAAGACATCTTCCTCTCTGGGACCCAGCAGTGCAGCGGAAACGAAAGAGG----- 235
DB 245 TCGTTCGCTCTTCCTGCTCGGAACGGGAGGAGTGGACACTTTTCTCCAAAGCAGCA 186
```

Qy	236	TTGACGAGGAGCGACGCGACATTTATCCAGAAGGATTTCTCTAGACGCTATT	295
Db	185	TCGAGGAAGAGCGCAAAATTTACCATGACATCATCCACAGGACACTACCAAGACACCTACT	126
Qy	296	ACAATCTGACCCCTGAAGACCATGATGGCATAGATGGTCCATCGCTTTTGTCTCTCAGG	355
Db	125	ACAACCTGACCATTAACCAACCTGATGGGTATGAATGGTGGCCACTATTGGCCACACG	66
Qy	356	CGCGGTTTCTGTGATGAACACAGACFCAGACATGTTTCATCAATGTTGACTATCTCACTGAA	414
Db	65	CTCTCTACTGTGATCAACACAGACAGACATGTTTGTTCACAGAGTATCHCAACAA	7

BE652985 574 bp mRNA linear EST 06-SEP-2000  
 UI-M-AM1-afz-e-01-0-UI.r1 NIH-BMAP\_MAM\_N Mus musculus cDNA clone  
 UI-M-AM1-afz-e-01-0-UI 5', mRNA sequence.  
 BE652985  
 BE652985.1 GI:9978848  
 EST.  
 house mouse.  
 Mus musculus  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 574)  
 Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genome Res. 6 (9), 791-806 (1996)  
 9704477  
 Contact: Chin, H  
 National Institute of Mental Health  
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
 20892-9643, USA  
 Tel: 301 443 1706  
 Fax: 301 443 9890  
 Email: mEST@mail.nih.gov  
 cDNA Library Preparation: M.B. Soares Lab Clone distribution:  
 Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It  
 should be noted that Bento Soares is generating a small number of  
 additional specialized non-redundant arrays of BMAP cDNAs whose  
 availability will be considered under appropriate and limited  
 collaborative arrangements  
 Seq primer: M13 Reverse.  
 Location/Qualifiers  
 1..574  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UI-M-AM1-afz-e-01-0-UI"  
 /clone\_lib="NIH\_BMAP\_MAM\_N"  
 /dev\_stage="27-32 days"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker; Site1: Not I; Site2: Eco RI; The  
 NIH\_BMAP\_MAM\_N library is a normalized library constructed  
 from mouse amygdala. The tag is a string of 5 nucleotides  
 present between the Not I site and the oligo-dT track.  
 The library was constructed as described by Bonaldo,  
 Lennon and Soares, Genome Research 6: 791-806, 1996.  
 Tissue provided by Ms. Annie Novakovich, Zivic-Miller  
 Laboratories."

BASE COUNT	161 a	113 c	121 g	178 t	1 others
ORIGIN					

	Query Match	13.7%	Score 114.8	DB 10	Length 574
Best Local Similarity	51.5%				
Pred. No. 9.3e-22					
Matches 291	Conservative	0	Mismatches 268	Indels	Gaps 1
QY	242	AGGAGGCCAGGACGGGGACATTCACGAAGGATTTCTAGACGCTTATTACAATC	301		
Db	569	AGAGAGGACAGATTCATGATTAATTCAGCAGGAATATTTTAAACAATACATAATC	510		

QY	302	TGACCCCTGAAGACCATGATGGGCATAGAATGGGTCCATCGCTTTTGTCTTCAGCGCGCGT	361
Db	509	TTGCCATTAAACAACCACTAATGGGTATGAATGGTGTGCAACATACTGTCACACATACTCCT	450
QY	362	TTGTGTGTAAGAACAGACTCGACACTGTTCATCAATGTTGACATATCTGACTGAACGTGTTTC	421
Db	449	ATGTTATGAAGAACGGACAGTAGCATGTTTGTCACACAGAATACTTTAATACACAAAGTTAC	390
QY	422	TGAAGAAAAACAGAACCAACCC-----GTTTTTTCACCTGGCTTCTTTGAAAACCTCAATGAGT	475
Db	389	TAAAGCCAGACCTGCCTCCTAGACATAACTATTTTACTGGCTATCTAATGAGAGGATATG	330
QY	476	TTCCCCATCAGCGAGCCATTACGCAAGTGGTTGTGTCAGTAAATCTGAATATCCGTGGGACA	535
Db	329	CACCGAACAGAAACAAGACACTAAGTGGTACATGCCACAGACCTTTTACCCAAGTGAGC	270
QY	536	GGTACCCACCATTCTGCTCCGGCACCGGTACGTGTTTTCTGGCGAGCTGGCGAGTGAGG	595
Db	269	GCTACCCCTGCTCTGCTCAGGAACCTGGTTATGTGTTTTCTGGGGATCTGGCAGAGAAGA	210
QY	596	TGTACAAATGTTCTCCAGAGGCTCCCATACATTAACCTGGAAGACGCTGTTGTGGGGCTCT	655
Db	209	TATTTAAGGTTTCTTTAAGGTATCGGTGCTTTTGCACCTTGAAGATGTATATGAGGATCT	150
QY	656	GCTCGAAGGCTGAACATCAGATGGAGGAGCTCCACTCCCGAGCGACCTTTTTCCTCAG	715
Db	149	GTCCTGCCAAGTTGAGAGTTGATCCTGTGCCCCCTCCCAATGAGTTCGTGTTCAATCACT	90
QY	716	GGGGCTTACGCTTCTCGGTATGCCTTTCAGGAGGATGCTGGCCTCCCACTTCATCAAGC	775
Db	89	GCGAGCTTCTTATTCAGCTGTAAATACAGCCACCTAATTACCTCTCATCAGTTCCCAAC	30
QY	776	CTCGGACTCTCTTGGACTACTGGCA	800
Db	29	CTAGTGAACCTGATGAATACTGGAA	5
RESULT	15		
BM548138			
LOCUS			
DEFINITION	BM548138	1101 bp	mRNA linear EST 20-FEB-2002
	AGENCOURT_6571527 NIH_MGC_124 Homo sapiens		cDNA clone IMAGE:573782
ACCESSION			5', mRNA sequence.
VERSION	BM548138		
KEYWORDS	BM548138.1	GI:18782503	
SOURCE			EST.
ORGANISM	Homo sapiens		human.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 1101)		
JOURNAL	NIH-MGC htp://mhc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs@mail.nih.gov		
	Tissue Procurement: Invitrogen		
	cDNA Library Preparation: Life Technologies, Inc.		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	htp://image.lnl.gov		
	Plate: LLM12738 Row: a Column: 15		

```

FEATURES
source
High quality sequence stop: 418.
Location/Qualifiers
1..1101
'organism="Homo sapiens"
'db_xref="taxon:9606"
'clone="IMAGE:5733782"
'clone_lib="NIH_MGC_124"
'tissue_type="hippocampus"

```

/lab\_host="DH10B"  
/note="Organ: brain; Vector: pCMV-SPORT6; Site\_1: EcoRV  
(destroyed); Site\_2: NotI; RNA source male hippocampus,  
age 27. Library is oligo-dT primed and directionally  
cloned (EcoRV site is destroyed upon cloning). Average  
insert size 1.4 kb, insert size range 0.9-4 kb. Library is  
normalized and enriched for full-length clones and was  
constructed by C. Gruber (Invitrogen). Research Genetics  
tracking code 012."  
BASE COUNT 304 a 239 c 238 g 319 t 1 others  
ORIGIN

Query Match 13.0%; Score 109.4; DB 13; Length 1101;  
Best Local Similarity 51.4%; Pred. No. 4.5e-20;  
Matches 285; Conservative 0; Mismatches 261; Indels 9; Gaps 1;  
Qy 20 TTACAAGAAAGAGCGGAGCTTCCTTAAGCTCCAGATACAGACTGCAGGCGAGACACCTC 79  
Db 222 TTACAGACAAGACTTTCACCTTCGAGAGCAATCAAACTGCTCATCAAAATC 281  
Qy 80 CTTCTCTCTCTCTGCTGCTGACCTCATCCACAACAGTTGGCTGAGGCGCATGGCCATCC 139  
Db 282 CATTTCTGCTCATTCGTGACCTCCACCTTCAGATGTGAAGCCAGGCGGCATTGA 341  
Qy 140 GCGACAGTGGGGGAAGAGAGAGATGGTGAAGGGAAGAGAGCTGAAGACATTTCTCTCC 199  
Db 342 GAGTTACTTGGGGTGAAGAAAGTCTTGGTGGGATATGAGGTTCTTACATTTTCTTAT 401  
Qy 200 TGGGACACACAGCAGTGCAGCGGGAACGAAGAGG-----TGGACAGGAGAGCC 250  
Db 402 TAGGCCAAGAGGCTGAAAAGGAAGACAAAATGTTGGCATTGTCCTTACAGGATGAACACC 461  
Qy 251 AGCGACAGGGGACATATCCAGAGGATTTCCAGAGGATTCCTAGACGCTATTACAATCTGACCCTGA 310  
Db 462 TTCCTTATGGTGACATAATCCGACAAGATTTTGTAGACATATAATAACCTGACCTTGA 521  
Qy 311 AGACCATGATGGGCATAGAGTGGTCCATCGCTTTGCTCAGCGCGCTTGTGTATGA 370  
Db 522 CAACCATATTGCGATTCAGGTGGGTAACTAGTGTTCGCCCAATGCCAAGTACGTAATGA 581  
Qy 371 AAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACCTGCTCTGAAGAAAA 430  
Db 582 AGACAGACACTGATGTTTTCATCAATCACTACTGGCAATTTAGTGAAGTATCTTTTAACCTAA 641  
Qy 431 ACAGAACACACAGGTTTTCACCTGGCTCTTGAAACTCAATGAGTTTCCCATCAGGCAGC 490  
Db 642 ACCACTCAGAGAAGTTTTCACAGGTTATCCTCTAATTGATAATTATTCCTATAGAGGAT 701  
Qy 491 CATTACAGCAAGTGGTTGTGAGTAATCTGAATATCGGTGGGACAGGTACCCACCAATCT 550  
Db 702 TTTACCAAAAACCATATTTCTTACCAGGAGTATCCTTTCAAGCGGTTCCTCCATACT 761  
Qy 551 GCTCGGCACCGGCT 565  
Db 762 GCAGGGGTAGGGGT 776

Search completed: April 11, 2003, 22:37:31  
Job time : 942.163 secs





Db 466 GTNTTYTNYNGNGTNGCNGNWSNCCNCCNGCNCARVNTYNGCNTAYGARWSN 525  
QY 250 CAGCGACAGCGGGGACATATCCAGAGGATTTCCCTAGACGCTCTATTACATCTGACCGCTG 309  
Db 526 MNGARTTYGAYGAYATHYNTCARTGGGAYTYACNGARGAYTYTYTYAAYTNACNYN 585  
QY 310 AAGACCATGATGGGCATAGAAATGGGTCATCGCTTTTGTGCTCAGCGGGCTTTGTGATG 369  
Db 586 AARGARYTNCARYTNCARMNTGGTNGTNGCNGNTGYCCNARGCNCAYTYTYATGYTN 645  
QY 370 AAAACAGACTCAGACATGTTTCATCAATGTGACTATCTGACTGAACTGCTTCTGAAGAAA 429  
Db 646 AARGNGAYGAYGAYGTNTTYTNCAYGTNCCNAAAYGTNYNGARTTYTYNGAYGNTGG 705  
QY 430 AACAGAACACAGGTTTTCAGCTGCTTCTGAACTCAATCAAGTTTCCCATCAGGCAG 489  
Db 706 GAYCCNGCNCARGAYTYTNYTNGTNGNGAYGTNATHMNGCNCNAYTYCCNCCNTAY 765  
QY 490 CCAITCAGCAAGTGGTTTGTCAAGTAACTGATATCCGTCGGGACAGGTCACCCACCATTC 549  
Db 766 ACNARGTNAARTATYTHCCNCCNWSNATGTAYMNGCNCNACNAYTYCCNCCNTAY 825  
QY 550 TGCTCCGCGCAGCGGTGCTGCTGCTGCGGAGCTGCGGAGTCAAGTGTACAAATGCTCTCC 609  
Db 826 CGCNGGNGGNGGNTAYGTNATGSHNMGNCNACNCTNMNGNVTNARGCNATHATG 885  
QY 510 AAGAGCGTCCCATACATTAACCTGGAAGACGCTGTTTGTGGGGCTGCTCCCTCGAAAGCTG 669  
-Db 886 GARGAYGNGARYTNYTNWSNATHGAYGAYGTNTTYTNGNATGTGYTINMNGWNYTN 945  
QY 670 AACATCAG 677  
Db 946 GGNYTNS 953

RESULT 2  
US-09-482-180A-1  
; Sequence 1, Application US/09482180A  
; Patent No. 6361985  
; GENERAL INFORMATION:  
; APPLICANT: Conklin, Darrell C.  
; APPLICANT: Yamamoto, Gayle  
; APPLICANT: Gao, Zeren  
; APPLICANT: Whitmore, Theodore B.  
; APPLICANT: Jaspers, Stephen  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING  
; TITLE OF INVENTION: BETA-1.3-GALACTOSYLTRANSFERASE HOMOLOG, ZNSP6  
; FILE REFERENCE: 98-80  
; CURRENT APPLICATION NUMBER: US/09/482,180A  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR FILING DATE: 1999-01-12  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 1420  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (135)...(1271)  
US-09-482-180A-1

Query Match 9.3%; Score 78.4; DB 4; Length 1420;  
Best Local Similarity 45.1%; Pred. No. 6.8e-15;  
Matches 299; Conservative 0; Mismatches 346; Indels 3; Gaps 1;  
QY 33 CGGGAACCTCTTAAGCTCCAGATACAGACTGCGAGGACAGACACCTCCCTCTCTCTCT 92  
Db 440 CGGAATTTCTCTATCTGCTGGAGCCTTCAGGCTTCCAGAGTACCTTTGCTCTCT 499  
QY 93 GCTGGTGACCTCATCCACAAACAGTTGGGTGAGCGCATGCCATCGCGCAGACGTGGG 152

Db 500 GGCCATCAAGTCACAGCTGTGTCAGTGAGGAGCGTGGGCTATCCGAGCAGCTGGGG 559  
QY 153 GAAAGAGAGG- --ATGCTGAAGGAAAGACAGCTGAAGACATTTCTCTCTCTCTGGGAGCCAC 209  
Db 560 CAGGGTGGGGGATGGCTAGGGCCGCGCAGCTGAAGCTGGTCTCTCTCTAGGGGTGGC 619  
QY 210 CAGCAGTGCAGCGGAAACGAAAGAGGTGGACGAGAGCCAGCCAGCAGCGGGACATTAT 269  
Db 620 AGGATCGCTCCCGCCAGCCAGCTGCTGGCCCTATGAGAGTAGGAGTTTGATGACATCCT 679  
QY 270 CCAGAAGGATTTCTCTAGACGCTCTATTACAATCTGACCTCTGAAGACCATGATGGGCATAGA 329  
Db 680 CCAGTGGGACTTCACTGAGGACTTCTTCAACCTGACGCTCAAGGAGCTGCACCTGCAGCG 739  
QY 330 ATGGGTCATCGCTTTTGTCTCAGCGGGCGCTTTGTGATGAAACAGACTCAGACATGTT 389  
Db 740 CTGGGTGGTGGCTGCTGCCCGCCAGGCCATTTATGCTAAGGAGATGAGCATGCTCT 799  
QY 390 CATCAATGTTGACTATCTGACTGAATGCTTCTGAAGAAAAACAGAAACACAGGTTTTT 449  
Db 800 TGTCCACGTCCTCAACGCTGTAGAGTTCTCTGATGGCTGGGACCCAGCCAGGACCTCCT 859  
QY 450 CACTGCTCTTGAAGTCAATGAGTTTCCCATCAGGAGCCATTCAGCAAGTGGTTGT 509  
Db 860 GGTGGAGATGTCTATCGCCCAAGCCCTGCCCCAACAGGAACACTAAGGTCAAATACTTCAT 919  
QY 510 CAGTAATCTGATATCCGTTGGGACAGTACCCACCATTTCTGCTCCGCGCAGCGCTACGT 569  
Db 920 CCCACCTCAATGTACAGGCCACCCACTACCCACCTATGCTGTGGGGAGGATATGT 979  
QY 570 GTTTCTGCGAGTGGGAGTCAAGTGTACAATGTCTCCAAAGAGCGTCCCATACATTAA 629  
Db 980 CATGTCAGGACAGCTGCGGCGCTCCAGGCTATCATGGAAGATGCTGAACCTCTCTC 1039  
QY 630 ACTGGAAGACGCTGTTTGTGGGCTCTGCTCGAAAGGCTGAACATCAG 677  
Db 1040 CATGTGATGATGCTTTTGTGGTATGTGCTTGAGGAGCGTGGGCTGAG 1087

RESULT 3  
US-09-459-133-3  
; Sequence 3, Application US/09459133  
; Patent No. 6416988  
; GENERAL INFORMATION:  
; APPLICANT: Conklin, Darrell C.  
; APPLICANT: Yamamoto, Gayle  
; APPLICANT: Jaspers, Stephen R.  
; APPLICANT: Gao, Zeren  
; TITLE OF INVENTION: BETA-1.3-GALACTOSYLTRANSFERASE HOMOLOGS  
; FILE REFERENCE: 98-77  
; CURRENT APPLICATION NUMBER: US/09/459,133  
; CURRENT FILING DATE: 1999-12-10  
; PRIOR FILING DATE: 1999-12-10  
; PRIOR FILING DATE: 1998-12-10  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 1191  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Degenerate sequence  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(1191)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-459-133-3

Query Match 8.2%; Score 68.6; DB 4; Length 1191;  
Best Local Similarity 28.3%; Pred. No. 8e-12;  
Matches 211; Conservative 95; Mismatches 428; Indels 12; Gaps 3;

QY	79	CCCTTCTCGTCGTGTAAGTTCAATCCACAAACAGTTGGCTGAGCGCATGCCCATC	138
Db		:    :    :    :    :    :    :    :    :    :    :    :    :	
Db	442	CNTAYTYNYTNVNGCNGTNAARWSNGARCCNGGNMNTTYGCNGARMGNCARGCNTN	501
QY	139	CGCGACAGCTGGGGAAAAGAGAGATTGGTGAAGGGAAGCAGTGTAAGACATTCTTCCTC	198
Db		:  :    :    :    :    :    :    :    :    :    :    :    :    :	
Db	502	MNGARACNTGGGWNWSCNCGNCCNGGNATHMGNTVNTNTTYTYTNNGWNSCCN	561
QY	199	CTGGGGACCACAGCAGTAGCGCGGGAAACGAAGAGGTGGACAGGAGACCGACGACAC	258
Db		:    :    :    :    :    :    :    :    :    :    :    :    :	
Db	562	GTINGGARCGNGNCCNGAYTYNGAWS---NYTNGTNGCNTGGARWSNMGMNGTAY	618
QY	259	GGGGACATTATCCAGAAGGATTTCCTAGAGCTATTACAATCTGACCCCTGAAGACCATG	318
Db		:  :    :    :    :    :    :    :    :    :    :    :    :    :	
Db	619	WSNGAYTYNYTNVTGGGAYTTYTYNGAYGTNCCNTTYAAYCARACNYTNAARGAYTN	678
QY	319	ATGGGCATAAGATGGGTCCATCGCTTTTCTCCACGGCGCGTTTGTGATGAAAAACAGAC	378
Db		:    :    :    :    :    :    :    :    :    :    :    :    :	
Db	679	YTYNYTNVNGCNTGYTYNGNMGNCAYTYGCCNACHGTNWSNTTYGTNYTMNGMCNCAR	738
QY	379	TCAGACATGTTCCATCAATGTTGACTATCTGACTGAACGTGT-----TCTGAAGAAAAAC	432
Db		:    :    :    :    :    :    :    :    :    :    :    :    :	
Db	739	GAYGAYGCNTTYGTNCAYACNCCNGCNYYTNVNGCNCAYYTNMGNGCNYTNCNCCNGCN	798
QY	433	AGAACACACCGAGTTTTTCTACTGGCTTCTTGAAACTCAATGAGTTTCCATCAGCGACCA	492
Db		:  :    :    :    :    :    :    :    :    :    :    :    :    :	
Db	799	WSNGCMGNWSNYTAYTYNGNGARGTNTTYACNCARGCNATGCCNYTMGNAAACCN	858
QY	493	TTCACCAAGTGGTTGTCCAGTAAATCTGAATATCCCTGGGACAGTAGTACCACCATCTGCG	552
Db		:    :    :    :    :    :    :    :    :    :    :    :    :	
Db	859	GGNGNCNTTYTAGTYNCCNGARWSNTTYTYTGARGNG--GNTAYCCNGCNTAYGCN	915
QY	553	TCCGSCACCGGTACGTGTTTCTTCGCGAGTGGCGAGTCAAGTGTACATGTCTCCAAG	612
Db		::   :	
Db	916	WSNGNGNGGNTATGNTNATHCGNGGNMGNYTNMGNCNCTGGYTYTNMGNGCNGCGCN	975
QY	613	AGCGTCCCATACATTAACCTGGAAGACGTGTTTGTGGGCTCTGCTCGAAGAGCTGAAC	672
Db		:      :      :      :      :      :      :      :      :      :	
Db	976	MNGTNGCNCCNTTYCCNTTYGARGAYGNTAYACMGNYTNTGYTHMGNGCYTNNGN	1035
QY	673	ATCAGATTGGAGGAGCTCCAGTCCGACCGACGCTTTTCCAGGGGCTTACGCTTCTCC	732
Db		:    :    :    :    :    :    :    :    :    :    :    :    :	
Db	1036	YTNCTNCCNCGNCAYCCNGNTTYTNACNGCNTGCCNCGNAYMGNACNGNGAY	1095
QY	733	GTATGCCCTTTCAGGAGGATCGTGCCCTGCCACTTCATCAAGCCTCGGACTCTCTTGGAC	792
Db		:    :    :    :    :    :    :    :    :    :    :    :    :	
Db	1096	CAYTYGCNTTYMGNAAYYTYTNVINGTNMGNCNYYTNNGNCCNARGCNWSNATHMGN	1155
QY	793	TACTGGGAGGCTCTAGAGAAATCCCG	818
Db		:      :      :      :      :      :      :      :      :      :	
Db	1156	YTNNGGAARCARYTNCARGAYCCNMG	1181

```

RESULT 4
US-09-459-133-14
; Sequence 14, Application US/09459133
; Patent No. 6416988
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Yamamoto, Gayle
; APPLICANT: Jaspers, Stephen R.
; APPLICANT: Gao, Zeren
; TITLE OF INVENTION: BETA-1,3-GALACTOSYLTRANSFERASE HOMOLOGS
; FILE REFERENCE: 98-77
; CURRENT APPLICATION NUMBER: US/09/459,133
; CURRENT FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/111,697
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 1167
; TYPE: DNA

```

RESULT 5  
US-09-085-097-2  
: Sequence 2, Application US/09055097  
: Patent No. 5955282  
: GENERAL INFORMATION:  
: APPLICANT: Hillman, Jennifer L.  
: APPLICANT: Guegler, Karl J.  
: APPLICANT: Corley, Neil C.  
: APPLICANT: Shah, Purvi  
: APPLICANT: Patterson, Chandra  
: TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR  
: NUMBER OF SEQUENCES: 7  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Incyte Pharmaceuticals, Inc.  
: STREET: 3174 Porter Drive  
: CITY: Palo Alto  
: STATE: California  
: COUNTRY: USA  
: ZIP: 94304



```

Best Local Similarity 48.5%; Pred. NO. 0.0096;
Matches 142; Conservative 0; Mismatches 148; Indels 3; Gaps 1;

QY 79 CCGTTCTCTGCTGCTGGTGACCTCATCCACAAACAGTTGGCTGAGCGCATGGCCATC 138
Db 524 CCTACTTGCTACTGGCTGTCAAATCAGAACAGAGACACTTTCCAGCAGCGGAGTGTG 583
QY 139 CGCGACAGCTGGGGGAAAGAGAGGATGGTGAAGGGAAAGCAGCTGAAGACATTTCTCTC 198
Db 584 AGSGAGACCTGGGGCAGCCCCAGTTGCTGGGACCCGGTTGCTCTTCTGCTGGGTCCCC 643
QY 199 CTGGGGACACACAGCAGCTGCAGCGGGAACGAAGAGGTGGACAGAGAGCGACCGACAC 258
Db 644 CTAGG---AATGGGGGGCGCTGACTTAGATCACTGGTGACGTGGGAAAGCGCGCTAT 700
QY 259 GGGGACATTATCAGAAAGGATTTCTTAGACGTCTATTACAATCTGACCCCTGAACACCATG 318
Db 701 GGTGACTACTGCTCTGGGACTTCTGTGATGTTCCCTACACCGGACACTCAAGGACCTG 760
QY 319 ATGGGCATGAAGTGGGTCCATCGCTTTTGTCTCAGGCGGCTTTGTGATGAA 371
Db 761 CTGCTGCTGACCTGGCTGAGCCACCACTGCCCGCATGTCAATTTTGTCTCTGCA 813

RESULT 9
US-09-404-650-12/c
; Sequence 12, Application US/09404650
; Patent No. 6309858
; GENERAL INFORMATION:
; APPLICANT: Dietrich, Paul S.
; APPLICANT: McGivern, Joseph G.
; TITLE OF INVENTION: T-TYPE CALCIUM CHANNEL VARIANTS; COMPOSITIONS THEREOF;
; TITLE OF INVENTION: AND USES
; FILE REFERENCE: R0043B-REG sequence listing
; CURRENT APPLICATION NUMBER: US/09/404,650
; CURRENT FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 6503
; TYPE: DNA
; ORGANISM: Rattus sp.
US-09-404-650-12

```

```

Query Match          4.5%; Score 37.4; DB 4; Length 6503;
Best Local Similarity 53.8%; Pred. No. 0.18;
Matches 77; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY      61  GACTGAGGCACACACTCCCTTCCTCGTCTCGTGTGTGACCTCATCCACAAACAGTTG 120
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      4197  GCGCGCACGACGTCAGATCGTTTGGTGATGTTTCGGGTGTCCACTCCCAACAATGGTAG 4138

QY      121  GCTGAGCGCATGCCATCCGGCAGACGCTGGGGGAAGAGAGAGATGTTGAAGGGAAGCAG 180
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      4137  AACTTTGCCCTTTGAAAAGCTGCACCCCGAGGATGCCGAAGATGATGAAGAAGGCACAGCAG 4078

QY      181  CTGAAGACATTCTTCTCCTCTGGG 203
      ||| ||| ||| ||| ||| |||
Db      4077  ATGAGCAGCATGTTCCCAATGGG 4055

RESULT 10
US-09-007-005-17
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; TITLE OF INVENTION: FUSIONS
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007, 005B

```

```

; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007, 005B

```

; CURRENT FILING DATE: 1998-01-14  
; EARLIER APPLICATION NUMBER: 60/035,963  
; EARLIER FILING DATE: 1997-01-27  
; EARLIER APPLICATION NUMBER: 60/064,491  
; EARLIER FILING DATE: 1997-11-06  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17  
; LENGTH: 289  
; TYPE: RNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Translation template  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(289)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-007-005-17

Query Match 4.4%; Score 36.6; DB 4; Length 289;  
Best Local Similarity 9.7%; Pred. No. 0.052;  
Matches 28; Conservative 109; Mismatches 152; Indels 0; Gaps 0;  
QY 149 GGGGAAAGAGAGATGCTGAAGGAAAGCAGCTGAAGACATTTCTCTCTGGGGACCA 208  
Db 1 RGRGRARCRARURURARURURARURURARURURARURURARURURARURURGRN 60  
QY 209 CCACGAGTGCAGCGGAAAGAGAGGTTGGACGAGGAGCCAGGACGAGGGGACATTA 268  
Db 61 RSRN 120  
QY 269 TCCAGAGGATTTCTAGACGCTTATTACAACTGACCTGAAGACCATGATGGGCATAG 328  
Db 121 RSRN 180  
QY 329 AATGGTCCATCGCTTTGCTCAGCGGCGTTTGTGATGAAACAGACATGACATGT 388  
Db 181 RSRN 240  
QY 389 TCATCAATGTTGACTATCTGACTGAACTGCTTCTGAGAGAAAACAGAAC 437  
Db 241 RCRURCRURGRGRCRUAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 289

RESULT 11  
US-09-244-796-17  
; Sequence 17, Application US/09244796  
; Patent No. 6281344  
; GENERAL INFORMATION:  
; APPLICANT: Szostak, Jack W.  
; APPLICANT: Roberts, Richard W.  
; APPLICANT: Liu, Rihe  
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN  
; FILE REFERENCE: 00786/350007  
; CURRENT APPLICATION NUMBER: US/09/244,796  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/035,963  
; EARLIER FILING DATE: 1997-01-27  
; EARLIER APPLICATION NUMBER: 60/064,491  
; EARLIER FILING DATE: 1997-11-06  
; EARLIER APPLICATION NUMBER: 09/007,005  
; EARLIER FILING DATE: 1998-01-14  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17  
; LENGTH: 289  
; TYPE: RNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Translation template  
; NAME/KEY: misc\_feature

; LOCATION: (1)...(289)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-244-796-17  
Query Match 4.4%; Score 36.6; DB 4; Length 289;  
Best Local Similarity 9.7%; Pred. No. 0.052;  
Matches 28; Conservative 109; Mismatches 152; Indels 0; Gaps 0;  
QY 149 GGGGAAAGAGAGATGCTGAAGGAAAGCAGCTGAAGACATTTCTCTCTGGGGACCA 208  
Db 1 RGRGRARCRARURURARURURARURURARURURARURURARURURARURURGRN 60  
QY 209 CCACGAGTGCAGCGGAAAGAGAGGTTGGACGAGGAGCCAGGACGAGGGGACATTA 268  
Db 61 RSRN 120  
QY 269 TCCAGAGGATTTCTAGACGCTTATTACAACTGACCTGAAGACCATGATGGGCATAG 328  
Db 121 RSRN 180  
QY 329 AATGGTCCATCGCTTTGCTCAGCGGCGTTTGTGATGAAACAGACATGACATGT 388  
Db 181 RSRN 240  
QY 389 TCATCAATGTTGACTATCTGACTGAACTGCTTCTGAGAGAAAACAGAAC 437  
Db 241 RCRURCRURGRGRCRUAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 289  
RESULT 12  
US-08-464-954A-1  
; Sequence 1, Application US/08464954A  
; Patent No. 6255069  
; GENERAL INFORMATION:  
; APPLICANT: BENOVIĆ, JEFFREY L.; GOMEZ, JORGE; KUNAPULI,  
; APPLICANT: PRIYA  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
; TITLE OF INVENTION: MODULATING THE ACTIVITY OF G PROTEIN-COUPLED RECEPTOR  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Jane Massey Licata, Esq.  
; STREET: Woodland Falls Corporate Park  
; STREET: 210 Lake Drive East, Suite 201  
; CITY: Cherry Hill  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 08002  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: WORDPERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/464,954A  
; FILING DATE: Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/076,084  
; FILING DATE: June 11, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jane Massey Licata  
; REGISTRATION NUMBER: 32,257  
; REFERENCE/DOCKET NUMBER: JEFF-0118  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (609) 779-2400  
; TELEFAX: (609) 779-8488  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2557  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: LINEAR

```

: ANTI-SENSE: NO
US-08-464-954A-1

Query Match 4.2%; Score 35; DB 4; Length 2557;
Best Local Similarity 48.7%; Pred. No. 0.61;
Matches 95; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 417 GCTTGTGAAGAAAACAGAACACCGAGTTTTCCTACCTGGCTTCTTGAAGAACTCAATGAGTT 476
Db 865 GCGCTTGGAGAAGAGAGGATCAAAAAGAGAGGAGGAGTCCATGGCCCTCAATGAGAA 924
QY 477 TCCCATCAGCCACCCATTACAGCAAGTGGTTGTCAAGTAATCTGAATCCGTGGGACAG 536
Db 925 CGAGATCCTCGAAGAGGTCAACAGTCAGTTGTGTCAACCTGGCTATSCCTACGAGAC 984
QY 537 GTACCCACCATCTGCTCGGCGCACCGCTACGTGTTTCTGSCGACGTGCGGAGTCAGGT 596
Db 985 CAACATGCACTGTGTTGTCTGACCATCATGAATGGGGGTGACGTGAAGTTCACAT 1044
QY 597 GTACAATGTCCTCAA 611
Db 1045 CTACAACATGGGCAA 1059

RESULT 13
US-09-007-005-32
: Sequence 32, Application US/09007005B
: Patent No. 6258558
: GENERAL INFORMATION:
: APPLICANT: Szostak, Jack W.
: APPLICANT: Roberts, Richard W.
: APPLICANT: Liu, Rihe
: TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
: FILE REFERENCE: 00786/350003
: CURRENT APPLICATION NUMBER: US/09/007,005B
: CURRENT FILING DATE: 1998-01-14
: EARLIER APPLICATION NUMBER: 60/035,963
: EARLIER FILING DATE: 1997-01-27
: EARLIER APPLICATION NUMBER: 60/064,491
: EARLIER FILING DATE: 1997-11-06
: NUMBER OF SEQ ID NOS: 33
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 32
: LENGTH: 248
: TYPE: RNA
: ORGANISM: Homo sapiens
US-09-007-005-32

Query Match 4.1%; Score 34.6; DB 4; Length 248;
Best Local Similarity 19.2%; Pred. No. 0.2;
Matches 41; Conservative 78; Mismatches 94; Indels 0; Gaps 0;

QY 145 ACGTGGGGAAGACAGAGGATGGTGAAGGGAAGACGCTGAAGACATTCCTCTCTGGG 204
Db 35 RCARARURARARARURGRGRCURGRARARGRARARARARARARARURGRA 94
QY 205 ACCACACGAGTCGCGGAAAGAGAGGTGGACACGAGGAGCCACGACCGGAC 264
Db 95 RURCURURGRARGRARGRARGRARCRURGRGRCURGRURARARARCRURJC 154
QY 265 ATTATCCAGAAGGATTCCTAGACGCTCTATTACAATCTGACCCCTGAAGACCATGATGGC 324
Db 155 RGRURARARCRARCRURGRARARARARARARARARARARARARURGRARARCRGHC 214
QY 325 ATAGAATGGGTCCATCGCTTTTGTCTCAGGG 357
Db 215 RURGRCRURARARCRURCRURGRRCRGR 247

RESULT 14
US-09-244-796-32
: Sequence 32, Application US/09244796

```

[illegible]

Search completed: April 11, 2003, 22:41:58  
Job time : 39.1953 secs

GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: April 11, 2003, 21:30:39 ; Search time 53.7426 Seconds  
(without alignments)  
13710.183 Million cell updates/sec

Title: US-09-914-152-2\_COPY\_492\_1331  
Perfect score: 840  
Sequence: 1 ttcaagaacagctcttctgt.....gggaagattgtcgccgtgc 840

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 593429 seqs, 438583890 residues  
Total number of hits satisfying chosen parameters: 1186858

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA.\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	130.4	15.5	1773	10	US-09-739-451-9
2	130.4	15.5	2095	9	US-09-992-598-208
3	130.4	15.5	2095	9	US-09-989-293A-208
4	130.4	15.5	2095	9	US-10-063-547-35
5	130.4	15.5	2095	9	US-09-989-735-208
6	130.4	15.5	2095	9	US-09-990-444-208
7	130.4	15.5	2095	9	US-09-989-730-208
8	130.4	15.5	2095	9	US-09-990-436-208
9	130.4	15.5	2095	9	US-09-991-181-208
10	130.4	15.5	2095	9	US-09-993-687-208
11	130.4	15.5	2095	9	US-09-989-734-208
12	130.4	15.5	2095	9	US-09-997-653-208
13	130.4	15.5	2095	9	US-10-174-590-161
14	130.4	15.5	2095	9	US-10-176-758-161
15	130.4	15.5	2095	9	US-10-063-616-35
16	130.4	15.5	2095	9	US-10-175-737-161
17	130.4	15.5	2095	9	US-09-993-667-208
18	130.4	15.5	2095	9	US-10-063-502-35
19	130.4	15.5	2095	9	US-10-173-706-161

20	130.4	15.5	2095	9	US-10-175-738-161	Sequence 161, App
21	130.4	15.5	2095	9	US-10-175-752-161	Sequence 161, App
22	130.4	15.5	2095	9	US-10-176-482-161	Sequence 161, App
23	130.4	15.5	2095	9	US-10-176-757-161	Sequence 161, App
24	130.4	15.5	2095	9	US-10-176-913-161	Sequence 161, App
25	130.4	15.5	2095	9	US-10-180-552-161	Sequence 161, App
26	130.4	15.5	2095	9	US-10-180-557-161	Sequence 161, App
27	130.4	15.5	2095	9	US-09-990-438-208	Sequence 208, App
28	130.4	15.5	2095	9	US-09-990-562-208	Sequence 208, App
29	130.4	15.5	2095	9	US-09-997-428-208	Sequence 208, App
30	130.4	15.5	2095	9	US-09-997-666-208	Sequence 208, App
31	130.4	15.5	2095	9	US-10-173-700-161	Sequence 161, App
32	130.4	15.5	2095	9	US-10-174-572-161	Sequence 161, App
33	130.4	15.5	2095	9	US-10-174-579-161	Sequence 161, App
34	130.4	15.5	2095	9	US-10-174-582-161	Sequence 161, App
35	130.4	15.5	2095	9	US-10-174-588-161	Sequence 161, App
36	130.4	15.5	2095	9	US-10-175-739-161	Sequence 161, App
37	130.4	15.5	2095	9	US-10-175-740-161	Sequence 161, App
38	130.4	15.5	2095	9	US-10-175-743-161	Sequence 161, App
39	130.4	15.5	2095	9	US-10-176-488-161	Sequence 161, App
40	130.4	15.5	2095	9	US-10-176-492-161	Sequence 161, App
41	130.4	15.5	2095	9	US-10-176-747-161	Sequence 161, App
42	130.4	15.5	2095	9	US-10-176-750-161	Sequence 161, App
43	130.4	15.5	2095	9	US-10-176-985-161	Sequence 161, App
44	130.4	15.5	2095	9	US-10-176-987-161	Sequence 161, App
45	130.4	15.5	2095	9	US-10-176-991-161	Sequence 161, App

ALIGNMENTS

RESULT 1  
US-09-739-451-9  
; Sequence 9, Application US/09739451  
; Patent No. US20010024813A1  
; GENERAL INFORMATION:  
; APPLICANT: Ebner, Reinhard  
; TITLE OF INVENTION: Dendriac and Brainiac-3  
; FILE REFERENCE: PF464  
; CURRENT APPLICATION NUMBER: US/09/739,451  
; PRIOR FILING DATE: 2000-12-19  
; PRIOR APPLICATION NUMBER: 09/213,364  
; PRIOR FILING DATE: 1998-12-17  
; PRIOR APPLICATION NUMBER: 60/077,687  
; PRIOR FILING DATE: 1998-03-12  
; PRIOR APPLICATION NUMBER: 60/108,928  
; PRIOR FILING DATE: 1998-11-17  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 1773  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-739-451-9

Query Match	15.5%	Score 130.4;	DB 10;	Length 1773;
Best Local Similarity	48.9%	Pred. No. 1.8e-33;		
Matches 387;	Conservative 0;	Mismatches 396;	Indels 9;	Gaps 1;
QY	20	TTTACAGAAAGACGGGAACCTTCCTTAAGCTCCAGATACAGACTGCAGGCAGACACCTC	79	
Db	157	TTTACAGACAGACTTTTCACCTTCACACTTCAGAGACATTCAAACTGCTCTCATCAAAATC	216	
QY	80	CCTTCCTCGCTCGCTGGTGACCTATCCCAACAAAGTTGGCTGAGCGCATGGCCATCC	139	
Db	217	CATTCTGTGTCATTCTGGTGACCTCCACCTTCAGATGTGAAGCCAGCGAGCCATT	276	
QY	140	GGCAGACGTGGGGAAGAGAGAGGATGGTGAAGGGAACAGCTGAAGACATTTCTCTCC	199	
Db	277	GAGTTACTTGGGGTGAAAAAAGTCTTGGTGGGATATGAGGTTCTTACATTTTCTTAT	336	
QY	200	TGGGGACCAACACAGCAGTGCAGCGGAAACGAAGAGG-----TGGACCAAGGAGCC	250	



Db 337 TAGCCAAAGAGCTGAAAGAGGAGCAAAATGTTGGCAATGTCTTAGAGGATGAACACC 396  
QY 251 AGCGACACGGGACATATCCAGAGGATTTCCCTAGACGCTCTATTACAATCTGACCCCTGA 310  
Db 397 TTCCTTTATGGTGACATAATCCGACAGATTTTGTAGACACATATAAATCACTGACCTTGA 456  
QY 311 AGACCATGATGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGGCGGCGTTTGTGATGA 370  
Db 457 AAACCATATGGCATTCAGGTGGGTAACTAGTTTCCCAATGCCAAGTAGTAATGA 516  
QY 371 AAACAGACTCAGACATGTTCAATGTTGACTATCTGACTGAACATGCTTCTGAAGAAAA 430  
Db 517 AGACAGACACTGATGTTTTCATCAATACTAGCAATTTAGTGAAGTATCTTTTAAACCTAA 576  
QY 431 ACAGAACCAACAGGTTTTTCACTGCTTCTTGAACCTCAATGAGTTTCCCATCAGGCAGC 490  
Db 577 ACCACTCAGAGAAGTTTTTTCACAGGTTATCCCTCTAATTGATAATTATTCCTATAGAGGAT 636  
QY 491 CATTGACGAAGTGTTCCTGAGTAATCTGAATATCCGTGGGACAGTACCCACCATCT 550  
Db 637 TTTACCAAAAAACCAATATTTCTTACCAGGATATCCCTTCAAGGTGTTCCCTCCCATACT 696  
QY 551 GCTCCGACACGGGTACGTGTTTCTGCGGAGTGGCGAGTGCAGGTGTACAATGTCTCCA 610  
Db 697 GCAGTGGGTTGGTTATATAATATGTCAGAGATTTGGTGCCAAGGATCTATGAATGATGG 756  
QY 611 AGACCGTCCCATACATTAATGAACTGGAAGGTTTGTGGGGTCTCTCCCTCGAAGGCTGA 670  
Db 757 GTCACGTTAAACCCCATCAAGTTTGAAGATGTTTATGTCGGGATCTGTTGCAATTTATTA 816  
QY 671 ACATCAGATGGAGGAGTCCACTCCACGCGGACCTTTTTCAGGGGCTTACGCTTCT 730  
Db 817 AAGTGAACATTCATATTCACAGACACAAATCTTTTCTTATATAGATCCATTTGG 876  
QY 731 CGTATGCCCTTCCAGGAGGATCGTGGCTGCCACTTCATCAAGCCTCGGACCTCTCTGG 790  
Db 877 ATGCTCTCAACTGAGAGTGTGATGTCAGGCCATGGCTTTTCTTCCAGGAGATCATCA 936  
QY 791 ACTACTGGCAGG 802  
Db 937 CTTTGGCAGG 948

RESULT 2  
US-09-992-598-208  
; Sequence 208, Application US/09992598  
; Patent No. US20020160384A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: P2730PIC20  
; CURRENT APPLICATION NUMBER: US/09/992,598  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088861



Db 1094 GTCAAGTAAACCCATCAAGTTTCAAGATGTTTATGTCGGGATCTGTTGAATTTATTTAA 1153  
QY 671 ACATCAGATTGGAGGAGCTCCACTCCAGCGAGCTTTTTTCCAGGGGCGTTACGGTCTTCT 730  
Db 1154 AACTGAACATTCATATTCCAGAGACACAAATCTTTCTTCTATATAGAAATCCATTTGG 1213  
QY 731 CCGTATGCCCTTCCAGAGAGATCGTGGCTGCCACTTCCATCAAGCCTCGGACTCTCTGG 790  
Db 1214 ATGCTGTCAACTGACAGCGTGTGATTGCAGCCCATGGCTTTCTTCCAAAGGAGATCATCA 1273  
QY 791 ACTACTGGCAGG 802  
Db 1274 CTTTGGCAGG 1285

RESULT 3

US-09-989-293A-208  
; Sequence 208, Application US/09989293A  
; Patent No. US2002017164A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kijavini, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; TITLE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: P2730PIC66  
; CURRENT APPLICATION NUMBER: US/09/989, 293A  
; CURRENT FILING DATE: 2001-11-20  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088876  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089105  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089440  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089512  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089514  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089532  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089598  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089599  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089600  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089653  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089907  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089908



RESULT 5  
 US-09-989-735-208  
 : Sequence 208, Application US/09989733  
 : Publication No. US20020193299A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Ashkenazi, Avi J.  
 : APPLICANT: Baker, Kevin P.  
 : APPLICANT: Botstein, David  
 : APPLICANT: Desnoyers, Luc  
 : APPLICANT: Eaton, Dan L.  
 : APPLICANT: Ferrara, Napoleone  
 : APPLICANT: Fong, Sherman  
 : APPLICANT: Gerber, Hanspeter  
 : APPLICANT: Gerritsen, Mary E.  
 : APPLICANT: Goddard, Audrey  
 : APPLICANT: Godowski, Paul J.  
 : APPLICANT: Grimaldi, J. Christopher  
 : APPLICANT: Gurney, Austin L.  
 : APPLICANT: Kijavlin, Ivar J.  
 : APPLICANT: Napier, Mary A.  
 : APPLICANT: Pan, James  
 : APPLICANT: Paoni, Nicholas F.  
 : APPLICANT: Roy, Margaret Ann  
 : APPLICANT: Stewart, Timothy A.  
 : APPLICANT: Tumas, Daniel  
 : APPLICANT: Watanabe, Colin K.  
 : APPLICANT: Williams, P. Mickey  
 : APPLICANT: Wood, William I.  
 : APPLICANT: Zhang, Zemin  
 : TITLE OF INVENTION: Secreted and Transmitted  
 : TITLE OF INVENTION: Acids Encoding  
 : FILE REFERENCE: P2730P1C61  
 : CURRENT APPLICATION NUMBER: US/09/989733  
 : CURRENT FILING DATE: 2001-11-19  
 : PRIOR APPLICATION NUMBER: 60/049787  
 : PRIOR FILING DATE: 1997-06-16  
 : PRIOR APPLICATION NUMBER: 60/062250  
 : PRIOR FILING DATE: 1997-10-17  
 : PRIOR APPLICATION NUMBER: 60/065186  
 : PRIOR FILING DATE: 1997-11-12  
 : PRIOR APPLICATION NUMBER: 60/065311  
 : PRIOR FILING DATE: 1997-11-13  
 : PRIOR APPLICATION NUMBER: 60/066770  
 : PRIOR FILING DATE: 1997-11-24  
 : PRIOR APPLICATION NUMBER: 60/075945  
 : PRIOR FILING DATE: 1998-02-25  
 : PRIOR APPLICATION NUMBER: 60/078910  
 : PRIOR FILING DATE: 1998-03-20  
 : PRIOR APPLICATION NUMBER: 60/083322  
 : PRIOR FILING DATE: 1998-04-28  
 : PRIOR APPLICATION NUMBER: 60/084600  
 : PRIOR FILING DATE: 1998-05-07  
 : PRIOR APPLICATION NUMBER: 60/087106  
 : PRIOR FILING DATE: 1998-05-28  
 : PRIOR APPLICATION NUMBER: 60/087607  
 : PRIOR FILING DATE: 1998-06-02  
 : PRIOR APPLICATION NUMBER: 60/087609  
 : PRIOR FILING DATE: 1998-06-02  
 : PRIOR APPLICATION NUMBER: 60/087759  
 : PRIOR FILING DATE: 1998-06-02  
 : PRIOR APPLICATION NUMBER: 60/087827  
 : PRIOR FILING DATE: 1998-06-03  
 : PRIOR APPLICATION NUMBER: 60/088021  
 : PRIOR FILING DATE: 1998-06-04  
 : PRIOR APPLICATION NUMBER: 60/088025  
 : PRIOR FILING DATE: 1998-06-04  
 : PRIOR APPLICATION NUMBER: 60/088026  
 : PRIOR FILING DATE: 1998-06-04  
 : PRIOR APPLICATION NUMBER: 60/088028  
 : PRIOR FILING DATE: 1998-06-04  
 : PRIOR APPLICATION NUMBER: 60/088029

; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22

; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 15.5%; Score 130.4; DB 9; Length 2095;
Best Local Similarity 48.9%; Pred. No. 2e-33;
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;

QY 20 TTTACAGAAGAGGGGAACCTCTTAAGTCCCGAGATACAGACTGAGCGAGACACCTC 79
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 494 TTTACAGAAGAGACTTTTCACTTCACTTCGAGAGCAATTCAACTGCTCTCATCAATC 553
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 80 CCTTCCTCGTCTGCTGCTGACCTATCCACAAACAGTTGGCTGAGCGCATGGCCATCC 139
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 554 CATTTCTGGTCATCTGCTGACCTCCACCTTCAGATGTGAAGCCAGCGAGCCATTA 613
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 140 GGCAGAGCTGGGGAAGAGAGAGGATGTTGAAGGGAACAGCACTGAAGACATTTCTCTCC 199
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 614 GAGTTACTTGGGTGAAAAAAGTCTTGGTGGGATATGAGTTCTTACATTTTCTTTAT 673  
Qy 200 TGGGACACACAGCAGTGCAGCGGAACGAAGAGG-----TGGACACAGGAGCC 250  
Db 674 TAGGCCAAGAGGCTGAAAAGGAAGACAAAATGTTGGCATTGTCTCTTAGAGGATGAACACC 733  
Qy 251 ACCGACACGGGACATATCCCAAGGATTTCTTAGAGGCTCTTATACAACTGACCCCTGA 310  
Db 734 TTCCTTATGGTGACATAATCCGACAGAGATTTTGTAGACACATATAAATACCTGACCTGA 793  
Qy 311 AGACCATGATGGGCATAGATGGTGCATCGCTTTTGTCTCAGCGCGGTTTGTGTATGA 370  
Db 794 AAACCATTTATGCAATTCAGTGGTGGTAACTGAGTTTCCGCCAAATGCCAAGTACGTAATGA 853  
Qy 371 AAACGACATCAGACATGTCATCAATGTTGACATATCTGACTGAACTGCTTCTGAAGAAA 430  
Db 854 AGACAGACACTGATGTTTTCATCAATACTGCAATTTAGTGAAGTATCTTTTAAACCTAA 913  
Qy 431 ACAGAACACACAGGTTTTCACCTGGCTTCTTGAACATCAATGAGTTTCCCATCAGGACG 490  
Db 914 ACCACTCAGAGAGTTTTCACAGGTTATGCTCTAATGTGATAATTTATCTTATAGAGGAT 973  
Qy 491 CATTGAGCAAGTGGTTGTGCTGATTAATGATATCGTGGGACAGGTACCCACCATCT 550  
Db 974 TTTACCAAAAACCCATATTTCTTACCAGGAGTATCTTTCAAGGTGTTCCCTCCATACT 1033  
Qy 551 GTCGCGCACCGGCTACGTTGTTTTCGGGACGCTGGCGAGTCAAGGTGATCAATGTCTCCA 610  
Db 1034 GCAGTGGGTTGGTTATATAATGTCAGAGATTTGGTGGCAAGGATCTATGAATGATGG 1093  
Qy 611 AGAGCGTCCCATCATTAACCTGGAAGACGTTGTTGGGGCTCTGCTGCGAAGGCTGA 670  
Db 1094 GTCAGTAAACCCCATCAAGTTTGAAGATGTTTATGTCGGGATCTGTTGAATTTATTA 1153  
Qy 671 ACATCAGATTGGAGGCTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCT 730  
Db 1.54 AAGTCAACATTCATATTCAGAGACACAAAATCTTTCTTCTATAGAAATCCATTGG 1213  
Qy 731 CCGTATGCTCTTCCAGAGGATCGTGGCTGCCACTTCATCAAGCCCGGACTCTCTTGG 790  
Db 1214 ATGCTGTCAACTGAGACGTGTGATTCAGAGCCCATGGCTTTCTTCCAGGAGATCA 1273  
Qy 791 ACTACTGGCAGG 802  
Db 1274 CTTTTGGCAGG 1285

RESULT 6  
US-09-990-444-208

; Sequence 208, Application US/09990444  
; Publication No. US20020193300A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnovers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kijavini, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.

; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; TITLE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: P2730P1C19  
; CURRENT APPLICATION NUMBER: US/09/990.444  
; PRIOR FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826

1	PRIOR FILING DATE: 1998-06-10	
2	PRIOR APPLICATION NUMBER: 60/008858	
3	PRIOR FILING DATE: 1998-06-11	
4	PRIOR APPLICATION NUMBER: 60/008861	
5	PRIOR FILING DATE: 1998-06-11	
6	PRIOR APPLICATION NUMBER: 60/008876	
7	PRIOR FILING DATE: 1998-06-11	
8	PRIOR APPLICATION NUMBER: 60/0089105	
9	PRIOR FILING DATE: 1998-06-12	
10	PRIOR APPLICATION NUMBER: 60/009440	
11	PRIOR FILING DATE: 1998-06-16	
12	PRIOR APPLICATION NUMBER: 60/0089512	
13	PRIOR FILING DATE: 1998-06-16	
14	PRIOR APPLICATION NUMBER: 60/0089514	
15	PRIOR FILING DATE: 1998-06-16	
16	PRIOR APPLICATION NUMBER: 60/0089532	
17	PRIOR FILING DATE: 1998-06-17	
18	PRIOR APPLICATION NUMBER: 60/0089538	
19	PRIOR FILING DATE: 1998-06-17	
20	PRIOR APPLICATION NUMBER: 60/0089598	
21	PRIOR FILING DATE: 1998-06-17	
22	PRIOR APPLICATION NUMBER: 60/0089599	
23	PRIOR FILING DATE: 1998-06-17	
24	PRIOR APPLICATION NUMBER: 60/0089600	
25	PRIOR FILING DATE: 1998-06-17	
26	PRIOR APPLICATION NUMBER: 60/0089653	
27	PRIOR FILING DATE: 1998-06-17	
28	PRIOR APPLICATION NUMBER: 60/0089801	
29	PRIOR FILING DATE: 1998-06-18	
30	PRIOR APPLICATION NUMBER: 60/0089907	
31	PRIOR FILING DATE: 1998-06-18	
32	PRIOR APPLICATION NUMBER: 60/0089908	
33	PRIOR FILING DATE: 1998-06-18	
34	PRIOR APPLICATION NUMBER: 60/0089947	
35	PRIOR FILING DATE: 1998-06-19	
36	PRIOR APPLICATION NUMBER: 60/0089948	
37	PRIOR FILING DATE: 1998-06-19	
38	PRIOR APPLICATION NUMBER: 60/0089952	
39	PRIOR FILING DATE: 1998-06-19	
40	PRIOR APPLICATION NUMBER: 60/0090246	
41	PRIOR FILING DATE: 1998-06-22	
42	PRIOR APPLICATION NUMBER: 60/0090252	
43	PRIOR FILING DATE: 1998-06-22	
44	PRIOR APPLICATION NUMBER: 60/0090254	
45	PRIOR FILING DATE: 1998-06-22	
46	PRIOR APPLICATION NUMBER: 60/0090429	
47	PRIOR FILING DATE: 1998-06-24	
48	PRIOR APPLICATION NUMBER: 60/0090349	
49	PRIOR FILING DATE: 1998-06-23	
50	PRIOR APPLICATION NUMBER: 60/0090355	
51	PRIOR FILING DATE: 1998-06-23	
52	PRIOR APPLICATION NUMBER: 60/0090429	
53	PRIOR FILING DATE: 1998-06-24	
54	PRIOR APPLICATION NUMBER: 60/0090431	
55	PRIOR FILING DATE: 1998-06-24	
56	PRIOR APPLICATION NUMBER: 60/0090445	
57	PRIOR FILING DATE: 1998-06-24	
58	PRIOR APPLICATION NUMBER: 60/0090472	
59	PRIOR FILING DATE: 1998-06-24	
60	PRIOR APPLICATION NUMBER: 60/0090535	
61	PRIOR FILING DATE: 1998-06-24	
62	PRIOR APPLICATION NUMBER: 60/0090540	
63	PRIOR FILING DATE: 1998-06-24	
64	PRIOR APPLICATION NUMBER: 60/0090542	
65	PRIOR FILING DATE: 1998-06-24	
66	PRIOR APPLICATION NUMBER: 60/0090557	
67	PRIOR FILING DATE: 1998-06-24	
68	PRIOR APPLICATION NUMBER: 60/0090676	
69	PRIOR FILING DATE: 1998-06-25	
70	PRIOR APPLICATION NUMBER: 60/0090678	
71	PRIOR FILING DATE: 1998-06-25	

?	PRIOR APPLICATION NUMBER: 60/090690
?	PRIOR FILING DATE: 1998-06-25
?	PRIOR APPLICATION NUMBER: 60/090694
?	PRIOR FILING DATE: 1998-06-25
?	PRIOR APPLICATION NUMBER: 60/090695
?	PRIOR FILING DATE: 1998-06-25
?	PRIOR APPLICATION NUMBER: 60/090696
?	PRIOR FILING DATE: 1998-06-25
?	PRIOR APPLICATION NUMBER: 60/090862
?	PRIOR FILING DATE: 1998-06-26
?	PRIOR APPLICATION NUMBER: 60/090863
?	PRIOR FILING DATE: 1998-06-26
?	PRIOR APPLICATION NUMBER: 60/091360
?	PRIOR FILING DATE: 1998-07-01
?	PRIOR APPLICATION NUMBER: 60/091478
?	PRIOR FILING DATE: 1998-07-02
?	PRIOR APPLICATION NUMBER: 60/091544
?	PRIOR FILING DATE: 1998-07-01
?	PRIOR APPLICATION NUMBER: 60/091519
?	PRIOR FILING DATE: 1998-07-02
?	PRIOR APPLICATION NUMBER: 60/091626
?	PRIOR FILING DATE: 1998-07-02
?	PRIOR APPLICATION NUMBER: 60/091633
?	PRIOR FILING DATE: 1998-07-02
?	PRIOR APPLICATION NUMBER: 60/091978
?	PRIOR FILING DATE: 1998-07-07
?	PRIOR APPLICATION NUMBER: 60/091982
?	PRIOR FILING DATE: 1998-07-07
?	PRIOR APPLICATION NUMBER: 60/092182
?	PRIOR FILING DATE: 1998-07-09

Query Match	15.5%	Score 130.4;	DB 9;	Length 2095;
Best Local Similarity	48.9%	Pred. No. 2e-33;		

Qy	20	TTTACAAGAAAGACGGGAAC	TCTCTTAAAGCTCCAGATACAGACTGCAGGACAGACACCTC	79	
Db	494	TTTACAGACAAGACTTTCACTT	CGAGACTTCAAAC	553	
Qy	80	CCTTCTCGTCTGCTGGTGA	CGCTCATCCACAAACAGTTGGCTGAGCGCATGGCCATCC	139	
Db	554	CATTTCTGGTGCACTTGGTGA	CGCTCCACCTTCAGATGTAAAGCCAGGACGGCCATTA	613	
Qy	140	GSCAGACGTGGGGAAAGAGAG	GATGGTGAAGGGAAGACAGCTGAAGACATCTTCCTCC	199	
Db	614	GAGTTACTTGGGGTGA	AAAAAAGCTTTGGTGGGATATGAGTCTTACATTTTTCTTAT	673	
Qy	200	TGGGACCACCAAGCAGCTG	CAGCGGAAACGAAAGG-----TGCACGAGGAGACC	250	
Db	674	TAGGCCAAGAGGCTGAA	AGGAGACAAATGTTGGCATTTGCTTTAGAGGATGAACACC	733	
Qy	251	AGCGACACGGGGACATTA	TCCAGAAGGATTTCTAGACGCTATTACAATCTGACCCCTGA	310	
Db	734	TTCTTTATGGTGACATA	TCCGACAGAATTTTTAGACACATATAATACCTGACCTGA	793	
Qy	311	AGACCATGATGGGCA	TAGAAATGGTCCATCGCTTTTGTCTCAGCGGGCGTTTGTGATGA	370	
Db	794	AAACCATTTAGGCATTT	CAGGTGGTAAC	TGAGTTTGGCCCAATGCCAAGTACGTAATGA	853
Qy	371	AAACAGACTCAGACATG	TTCAATCTGATCTGACTGAAC	TGCTTCTGAAGAAA	430
Db	854	AGACAGACACTGATG	TTTTTCATCACTAC	TGGCAATTTAGTGAAGTATCTTTTAAACCTAA	913
Qy	431	ACAGAACAACACAGG	TTTTTCAGTGGCTTCTTGAAC	CTCAATGAGTTTCCCATCAGGCAGC	490
Db	914	ACCAC	TGAGAAGTTTTTCAGAGTTATCCTCTAAATGATAATTA	TCTCTATAGAGAT	973
Qy	491	CATTACAGCAAGTGG	TTGTGCAGTAATCTGAATATCCGTGGGACAGGTACCCACCATCT	550	
Db	974	TTTACCAAAAAACCA	TATTTCTTACAGGAGTATCCTTTCAAGGTGTTCCCTCCCATACT	1033	
Qy	551	GCTCCGGCACCGGTAC	GTGTTTTCTGGCGAGGTGGCGAGTCAAGGTGACAAATGTC	TCCCA	610



Db 1034 GCAGTGGTGGTTATATATATCCAGAGATTTGGTGCCCAAGGATCTATGAATGATGG 1093  
QY 611 AGAGCGTCCCATACATAAAGTGAAGACGCTGTTGTGGGCTCTGCGCTCGAAGGCTGA 670  
Db 1094 GTCAACGTAACCCATCAAGCTTGAAGATGTTTATGCGGATCTGTTTGAATTTATTA 1153  
QY 671 ACATCAGATGGAGAGCTCCACTCCAGCGGACCTTTTTCCAGGGGGCTTACGCTTCT 730  
Db 1154 AAGTGAACATTCATATTCAGAGACACAAATCTTTCTTCTATATAGAATCCATTTGG 1213  
QY 731 CGATGCTCTTCAGGAGGATCGTGCGCTGCCACTTCATCAAGCCTCGGACTCTCTTGG 790  
Db 1214 ATGCTGTCACTGAGACGCTGATTGACGCCATGGCTTTCTTCCAGGAGATCATCA 1273  
QY 791 ACTACTGGCAGG 802  
Db 1274 CTTTTGGCAGG 1285  
RESULT 7  
US-09-989-730-208  
; Sequence 208, Application US/09989730  
; Publication No. US20020197674A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2730PIC69  
; CURRENT APPLICATION NUMBER: US/09/989,730  
; CURRENT FILING DATE: 2001-11-20  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088876  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089105  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089440  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089512  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089514  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089532  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089598  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089599  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089600  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089653  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089801

; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07

; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 15.5%; Score 130.4; DB 9; Length 2095;
Best Local Similarity 48.9%; Pred. No. 2e-33;
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;

Qy 20 TTTACAGAAAGACGGGAAGCTTCCTTAAGCTCCAGATACAGACTGCAGGACAGACACCTC 79
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 494 TTTACAGACAAAGACTTTTCACTTCACACTTCAGAGCATTCAAACTGCTCTCATCAAATC 553
Qy 80 CTTTCCTCGTCTGCTGACCTCATCCCAACAACAGTTGGCTGAGGCGATGGCATCC 139
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 554 CATTTCTGGTCATCTCTGGTGCACCTCCCACTTCAGATGTGAAGCCAGGCGCATTA 613
Qy 140 GGCAGACGTGGGGGAAAGAGAGGATGGTGAAGGAAAGACGCTGAAGACATTTCTCTCC 199
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 614 GAGTTACTTGGGTGAAAAAAGTCTTGGTGGGATATGAGGTTCTTACATTTTCTTAT 673
Qy 200 TGGGGACACACGACGTCAGCGGAAAGAGAGG-----TGGACACGAGGAGCC 250
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 674 TAGGCCAAGAGGCTGAAAGGAAACACAAATGTTGGCATTTGCTTAGAGGATGAACAC 733
Qy 251 AGCGACACGGGGACATTATCCAGAAGGATTTCTTAGAGTCTATTACAACTGACCTCA 310
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 734 TTCCTTATGGTGACATAATCCGACAGATTTTTTAGACACATATAATAACCTGACCTTGA 793
Qy 311 AGACCATGATGGGCATAGATGGTCCATCGCTTTTGTCTCAGCGCGGCTTTGTGATCA 370
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 794 AAACCATTTATGGCATTTCAGTGGTACTAGTTTGTCCCAATGCCCAAGTACGTAATGA 853
Qy 371 AAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAAGTCTTCTGAAGAAA 430
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 854 AGACAGACACTGATGTTTTCATCAATGTTGCAATTTAGTCAAGTATCTTTTAAACCTAA 913
Qy 431 ACAGACACACGAGTTTTCACACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGACG 490
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 914 ACCACTCAGAGAAAGTTTTCACAGGTTATCTCTAATTGATAATTATTCCTATAGAGGAT 973
Qy 491 CATTCAGCAAGTGGTTTGTCAAGTAAATCTGAATATCGTGGGACAGGTACCCACCATCT 550
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 974 TTTACAAAAAACCCATATTTCTTACAGGATATCTCTTCAAGGTGTTCCCTCCATACT 1033
Qy 551 GTCGGGACACCGCTACGTGTTTCTGGCGACGTGGCGAGTCAGGTGTACATGTCTCCA 610
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1034 GCAGTGGGTGGTTATATAATCTCCAGAGATTTGGTGCCNAGGATCTATGAATGATGG 1093
Qy 611 AGAGCGTCCATACATTAACCTGGAAGACGTGTTTGTGGGCTCTGCCTCGAAAGGCTGA 670
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1094 GTCACGTAAACCCCATCAAGTTTGAAGATGTTTATGTCGGGATCTGTTGATTTATTA 1153
Qy 671 ACATCAGATTGGAGAGCTCCACTCCAGCGGACCTTTTTTCCAGGGGGCTTTACGCTTCT 730
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1154 AAGTGAACATTCATATTCAGAGACACAAATCTTTTCTTCTATATAGAATCCATTTGG 1213
Qy 731 CCGTATGCTCTTCAGGAGGATCGTGGCTGCCACTTCATCAAGCCTCGGACTCTCTGG 790
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1214 ATGCTGTCAACTGAGACGTGATTTCAGCCCATGCTTTTCTTCCAAAGAGGATCATCA 1273
Qy 791 ACTACTGGCAGG 802
||||| |||||
Db 1274 CTTTTTGGCAGG 1285

RESULT 8
US-09-990-436-208
; Sequence 208, Application US/09990436
; Publication No. US20020198148A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730PIC14  
CURRENT APPLICATION NUMBER: US/09/990,436  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-08-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-08-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-08-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-08-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-08-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-08-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-08-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-08-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-08-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-08-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-08-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-08-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-08-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-08-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444

794 AAACCAATTATGGCAATTCAGTGGGTAACTGAGTTTTGCCCAATGCCAAGTACGTAATGA 853  
QY 371 AAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACGTCTTCTGAAGAAA 430  
Db 854 AGACAGACACTGATGTTTTCATCAATGTTGACTGAACGTCTTCTGAAGAAA 913  
QY 431 ACAGAACAAACAGGTTTTTCACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGCAGC 490  
Db 914 ACCACTCAGAGAAAGTTTTTTCACAGGTTATCTCTAATTGATAATTTATCCATAGAGGAT 973  
QY 491 CATTGAGCAAGTGGTTTTCAGTAAATCTGAATATCCGTGGGAGCAGGTTACCCACCATCT 550  
Db 974 TTACCAAAAAACCCCATATTTTACAGGAGTATCTCTTCAAGGTGTTCCCTCCATCT 1033  
QY 551 GTCCTGGCAGCGCTACGTGTTTCTGGCAGCTGGGCGAGTCAGGTGTACATGTCTCCA 610  
Db 1034 GCAGTGGGTGGTTTATATAATGTCAGAGATTTGGTCCCAAGGATCTATGAAATGATGG 1093  
QY 611 AGAGCGTCCCATACATTAACCTGAAGACGTGTTTGGGGGCTCTGCTCGAAAGGCTGA 670  
Db 1094 GTCAGGTAACCCCATCAAGTTTGAAGATGTTTATGTCGGGATCTGTTTGAATTTATTA 1153  
QY 671 ACATCAGATTGGAGGAGCTCCACTCCAGCGGACCTTTTCCAGGGGGCTTACGCTTCT 730  
Db 1154 AAGTGAACATTCATATTCAGAGAGACACAATCTTTTCTTCTATATAGAATCCATTTGG 1213  
QY 731 CGGTATGCTCTTCAGGAGGATCGTGGCCTGGCCTGCACTTCATCAAGCCTCGGACTCTCTGG 790  
Db 1214 ATGCTGTCACTGAGACGTGTGATTGCGAGCCCATGCTTTTCTTCCAGGAGATCATCA 1273  
QY 791 ACTACTGGCAGG 802  
Db 1274 CTTTTGGCAGG 1285

## RESULT 9

US-09-991-181-208

; Sequence 208, Application US/09991181

; Publication No. US20020197615A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Guiney, Austin L.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Napier, Mary A.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: P2730P1C53

; CURRENT APPLICATION NUMBER: US/09/991,181

; CURRENT FILING DATE: 2001-11-16

; PRIOR APPLICATION NUMBER: 60/049787

; PRIOR FILING DATE: 1997-06-16

; PRIOR APPLICATION NUMBER: 60/062250

Query Match 15.5%; Score 130.4; DB 9; Length 2095;  
Best Local Similarity 48.9%; Pred. No. 2e-33;  
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;  
QY 20 TTACAGAAAGCGGAACTTCTTAAGCTCCAGATACAGACTGCAGGCGAGCACACCTC 79  
Db 494 TTACAGAAAGACTTTCACTTCCACTTCGAGAGCAATTCAACTGCTCTCATCAAAATC 553  
QY 80 CCTTCCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 139  
Db 554 CATTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 613  
QY 140 GGCAGACGTGGGGAAGAGGATGCTGAAGGGAAGACAGCTGAAGACATTTCTTCTCC 199  
Db 614 GAGTTACTTGGGTTGAAAAGAAAGCTTCTGGTGGGATATGAGGTTCTTACATTTCTTAT 673  
QY 200 TGGGGACCAACAGCAGTGCAGCGGAACAGAAAGAGG-----TGGCAGGAGAGGC 250  
Db 674 TAGGCCAAGAGGCTGAAAAGGAAGACAAAATGTTGGCATTCCTTAGAGGATGAACACC 733  
QY 251 ACCGACACGGGACATATATCCAGAGAGTTCCTTAGAGCTGCTTACAAATCTGACCCCTGA 310  
Db 734 TTTCTTTATGCTGACATAATCCGACAGATTTTTCAGACACATATAAATTAACCTGACCTTGA 793  
QY 311 AGACCATGATGGCATAGATGGTCCATCGCTTTTGTCTCCTCAGCGCGCTTTGTGATGA 370



; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 15.5%; Score 130.4; DB 9; Length 2095;
Best Local Similarity 48.9%; Pred. No. 2e-33;
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;

QY 20 TTATCAAGAAAGACGGGAATCTTCTTAAGTCCAGATACAGACTGCAGGCGACACCTTC 79
Db 494 TTATCAAGAAAGACCTTCACTTCACTTCAGAGCAATTCACACTGCTCTCATCAAAATC 553
QY 80 CCTTCCTCGTCTGCTGGTCACTCATCCACAAACAGTTGGCTGAGGCGCATGGCATCC 139
Db 554 CATTTCTGGTCACTTCTGGTCACTCCACCTTCAGATGTGAAAGCCAGGCGACATTA 613
QY 140 GGCAGACGTGGGGAAGAGAGAGATGTGAAGGAAAGACAGTGAAGACATTTCTCTCC 199
Db 614 GAGTTACTTGGGTGAAAAAAGTCTTGGTGGGATATGAGTCTTACATTTTCTTAT 673
QY 200 TGGGACACACAGCAGTGCAGCGGAAACGAAGAGG-----TGGACACAGGAGGCC 250
Db 674 TAGGCCAAGAGGCTGAAAGAGAGACAAATGTTGGCAATTCCTTAGAGGATGACACC 733
QY 251 AGGCACACGGGACATATATCCGAAGATTTCCCTAGAGCTTATTACAACTGACCCCTGA 310
Db 734 TTCTTTATGGTGACATAAATCCGACAAAGATTTTATAGACACATAAATAACCTGACCTGA 793
QY 311 AGACCATGATGGCATAGATAGTGGTCCATCGCTTTTGTCTCCCTCAGCGCGGTTTGTGATGA 370
Db 794 AAACCATATGGCATCAGTGGGTAACTGAGTTTGGCCCAATGGCCAAAGTACGTAATGA 853
QY 371 AAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAATGCTTCTGAAGAAA 430
Db 854 AGACACACACTGATGTTTTCATCAATGCTGCAATTTAGTGAAGTATCTTTTAAACCTAA 913
QY 431 ACAGAAACACCGGTTTTCCTACCTGGCTCTTGAACACTCAATGAGTTTCCCATCAGGCGAGC 490
Db 914 ACCACTCAGAGAGATTTTTCACAGGTTATCCTCTAATTGATAATTATTCCTATAGAGAT 973
QY 491 CATTCAGCAAGTGGTTTGTGTCAGTAAATCTGAATATCCGTGGGAGAGGTACCCACCATCT 550
Db 974 TTACCAAAAACCCCATATTTCTTACAGAGAGTATCTCTTCAAGGTGTCTCCCTCCATCT 1033
QY 551 GCTCGGCGACCGGCTACGTGTTTCTTGGCGAGGTGGCGAGTCAAGTGTGACATGCTCTCCA 610
Db 1034 GCAGTGGGTGGTATATAAATGTCAGAGATTTGGTGGCAAGGATCTATGAAATGATGG 1093
QY 611 AGAGGCTCCCATACATTAACCTGGAAGACGTGTTTGTGGGCTCTGCTCGAAAGGCTGA 670
Db 1094 GTCACGTAAACCCCATCAAGTTTGAAGATCTTTATGTCGGATCTGTTGATTTATTAA 1153
QY 671 ACATCAGATTGAGGAGCTCCACTCCAGCCGACCTTTTTCAGGGGGCTTACGCTTCT 730
Db 1154 AAGTGAACATTTCATATCCAGAGAGACAAATCTTTTCTTCTATATAGAAATCCATTGG 1213
QY 731 CGGTATGCTCTTCAGGAGGATCGTGGCGTGGCCACTTCATCAAGCCCTCGGACTCTCTTGG 790

Db 1214 ATGTCTGTCAACTGAGACGTGTGATTGCAGCCCATGCTTTTCTTCCAGGAGATCATCA 1273
QY 791 ACTACTGGCAGG 802
Db 1274 CTTTITGGCAGG 1285

RESULT 10
US-09-993-687-208
; Sequence 208, Application US/09993687
; Publication No. US20020198149A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferfara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Geiritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C11
; CURRENT APPLICATION NUMBER: US/09/993,687
; CURRENT FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025

;	PRIOR FILING DATE:	1998-06-04
;	PRIOR APPLICATION NUMBER:	60/088026
;	PRIOR FILING DATE:	1998-06-04
;	PRIOR APPLICATION NUMBER:	60/088028
;	PRIOR FILING DATE:	1998-06-04
;	PRIOR APPLICATION NUMBER:	60/088029
;	PRIOR FILING DATE:	1998-06-04
;	PRIOR APPLICATION NUMBER:	60/088030
;	PRIOR FILING DATE:	1998-06-04
;	PRIOR APPLICATION NUMBER:	60/088033
;	PRIOR FILING DATE:	1998-06-04
;	PRIOR APPLICATION NUMBER:	60/088326
;	PRIOR FILING DATE:	1998-06-04
;	PRIOR APPLICATION NUMBER:	60/088167
;	PRIOR FILING DATE:	1998-06-05
;	PRIOR APPLICATION NUMBER:	60/088202
;	PRIOR FILING DATE:	1998-06-05
;	PRIOR APPLICATION NUMBER:	60/088212
;	PRIOR FILING DATE:	1998-06-05
;	PRIOR APPLICATION NUMBER:	60/088217
;	PRIOR FILING DATE:	1998-06-05
;	PRIOR APPLICATION NUMBER:	60/088655
;	PRIOR FILING DATE:	1998-06-09
;	PRIOR APPLICATION NUMBER:	60/088734
;	PRIOR FILING DATE:	1998-06-10
;	PRIOR APPLICATION NUMBER:	60/088738
;	PRIOR FILING DATE:	1998-06-10
;	PRIOR APPLICATION NUMBER:	60/088742
;	PRIOR FILING DATE:	1998-06-10
;	PRIOR APPLICATION NUMBER:	60/088810
;	PRIOR FILING DATE:	1998-06-10
;	PRIOR APPLICATION NUMBER:	60/088824
;	PRIOR FILING DATE:	1998-06-10
;	PRIOR APPLICATION NUMBER:	60/088826
;	PRIOR FILING DATE:	1998-06-10
;	PRIOR APPLICATION NUMBER:	60/088858
;	PRIOR FILING DATE:	1998-06-11
;	PRIOR APPLICATION NUMBER:	60/088861
;	PRIOR FILING DATE:	1998-06-11
;	PRIOR APPLICATION NUMBER:	60/088876
;	PRIOR FILING DATE:	1998-06-11
;	PRIOR APPLICATION NUMBER:	60/089105
;	PRIOR FILING DATE:	1998-06-12
;	PRIOR APPLICATION NUMBER:	60/089440
;	PRIOR FILING DATE:	1998-06-16
;	PRIOR APPLICATION NUMBER:	60/089512
;	PRIOR FILING DATE:	1998-06-16
;	PRIOR APPLICATION NUMBER:	60/089514
;	PRIOR FILING DATE:	1998-06-16
;	PRIOR APPLICATION NUMBER:	60/089532
;	PRIOR FILING DATE:	1998-06-17
;	PRIOR APPLICATION NUMBER:	60/089538
;	PRIOR FILING DATE:	1998-06-17
;	PRIOR APPLICATION NUMBER:	60/089598
;	PRIOR FILING DATE:	1998-06-17
;	PRIOR APPLICATION NUMBER:	60/089599
;	PRIOR FILING DATE:	1998-06-17
;	PRIOR APPLICATION NUMBER:	60/089600
;	PRIOR FILING DATE:	1998-06-17
;	PRIOR APPLICATION NUMBER:	60/089653
;	PRIOR FILING DATE:	1998-06-17
;	PRIOR APPLICATION NUMBER:	60/089801
;	PRIOR FILING DATE:	1998-06-18
;	PRIOR APPLICATION NUMBER:	60/089907
;	PRIOR FILING DATE:	1998-06-18
;	PRIOR APPLICATION NUMBER:	60/089908
;	PRIOR FILING DATE:	1998-06-18
;	PRIOR APPLICATION NUMBER:	60/089947
;	PRIOR FILING DATE:	1998-06-19
;	PRIOR APPLICATION NUMBER:	60/089948
;	PRIOR FILING DATE:	1998-06-19
;	PRIOR APPLICATION NUMBER:	60/089952
;	PRIOR FILING DATE:	1998-06-19

7	PRIOR APPLICATION NUMBER: 60/090244
7	PRIOR FILING DATE: 1998-06-22
7	PRIOR APPLICATION NUMBER: 60/090252
7	PRIOR FILING DATE: 1998-06-22
7	PRIOR APPLICATION NUMBER: 60/090254
7	PRIOR FILING DATE: 1998-06-22
7	PRIOR APPLICATION NUMBER: 60/090349
7	PRIOR FILING DATE: 1998-06-23
7	PRIOR APPLICATION NUMBER: 60/090355
7	PRIOR FILING DATE: 1998-06-23
7	PRIOR APPLICATION NUMBER: 60/090429
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090431
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090435
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090444
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090445
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090472
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090535
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090540
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090542
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090557
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090676
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090678
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090690
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090694
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090695
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090696
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090862
7	PRIOR FILING DATE: 1998-06-26
7	PRIOR APPLICATION NUMBER: 60/090863
7	PRIOR FILING DATE: 1998-06-26
7	PRIOR APPLICATION NUMBER: 60/091360
7	PRIOR FILING DATE: 1998-07-01
7	PRIOR APPLICATION NUMBER: 60/091478
7	PRIOR FILING DATE: 1998-07-02
7	PRIOR APPLICATION NUMBER: 60/091544
7	PRIOR FILING DATE: 1998-07-01
7	PRIOR APPLICATION NUMBER: 60/091519
7	PRIOR FILING DATE: 1998-07-02
7	PRIOR APPLICATION NUMBER: 60/091626
7	PRIOR FILING DATE: 1998-07-02
7	PRIOR APPLICATION NUMBER: 60/091633
7	PRIOR FILING DATE: 1998-07-02
7	PRIOR APPLICATION NUMBER: 60/091978
7	PRIOR FILING DATE: 1998-07-07
7	PRIOR APPLICATION NUMBER: 60/091982
7	PRIOR FILING DATE: 1998-07-07
7	PRIOR APPLICATION NUMBER: 60/092182
7	PRIOR FILING DATE: 1998-07-09

Query Match	15.5%	Score 130.4;	DB 9;	Length 2095;
Best Local Similarity	48.9%	Pred. No. 2e-33;		

Matches	387	Conservative	0	Mismatches	396	Indels	9	Gaps	1
---------	-----	--------------	---	------------	-----	--------	---	------	---

**Qy** 20 TTTACAAGAAACGCGGAATTCCTTAAGCTCCAGATACAGACTGCAGGCCAGACACCTC 79  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
**Dd** 494 TTTCACAGCAAGACTTTCACCTCACACTTCGAGAGCATTCAAACTGTCTCATCAAANTC 553





7	PRIOR APPLICATION NUMBER: 60/090557
7	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090676
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090678
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090690
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090694
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090695
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090696
7	PRIOR FILING DATE: 1998-06-25
7	PRIOR APPLICATION NUMBER: 60/090862
7	PRIOR FILING DATE: 1998-06-26
7	PRIOR APPLICATION NUMBER: 60/090863
7	PRIOR FILING DATE: 1998-06-26
7	PRIOR APPLICATION NUMBER: 60/091360
7	PRIOR FILING DATE: 1998-07-01
7	PRIOR APPLICATION NUMBER: 60/091478
7	PRIOR FILING DATE: 1998-07-02
7	PRIOR APPLICATION NUMBER: 60/091544
7	PRIOR FILING DATE: 1998-07-01
7	PRIOR APPLICATION NUMBER: 60/091519
7	PRIOR FILING DATE: 1998-07-02
7	PRIOR APPLICATION NUMBER: 60/091626
7	PRIOR FILING DATE: 1998-07-02
7	PRIOR APPLICATION NUMBER: 60/091633
7	PRIOR FILING DATE: 1998-07-02
7	PRIOR APPLICATION NUMBER: 60/091878
7	PRIOR FILING DATE: 1998-07-07
7	PRIOR APPLICATION NUMBER: 60/091982
7	PRIOR FILING DATE: 1998-07-07
7	PRIOR APPLICATION NUMBER: 60/092182
7	PRIOR FILING DATE: 1998-07-09

QY	20	TTTACAGAAAGACGGGAACCTTCCTTAAAGTCCAGATACAGACTCGACGACAGACCTC	79
Db	494	TTTACAGACAAGACTTTCACTTCACACTTCGAGAGCATTTCAAAGTCTCTCATCAAATC	553
QY	80	CCTTCCCTGCTCTGCTGACCTCATCCACAAACAGTTGGCTGAGCGCATGCCATCC	139
Db	554	CAITTCGGTCAATCTCGTGAACCTCCACCCCTTCAGATGTGAAGCCAGCGAGCCATTA	613
QY	140	GGCAGACTGGGGAAAGAGAGAGTGGTGAAGGAAAGCAGCTCAAGACATTTCTTCCTCC	199
Db	614	GAGTTACTTGGGGTGA AAAAAGCTTGTGGTGGGATAGAGGTTCTTACATTTTCTTAT	673
QY	200	TGGGACCCACGACGTGACGGGAACGAAAGG-----TGGACCAGAGAGCC	250
Db	674	TAGCCCAAGAGGCTGAAAAGGAGACAAAATGTGGCATTTGCTTAGAGGATGAACACC	733
QY	251	AGGCACACGGGGACATATATCCAGAAAGATTTCCTAGACGCTCTATTACAATCTCACCCCTGA	310
Db	734	TTCTTTATGTGACATAATCCGACAAGATTTTTAGACACATATAATAACCTGACCTTGA	793
QY	311	AGACCATGATGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGGGGGCGTTTGTGATGA	370
Db	794	AAACCATATTGGCATTCAGGTGGGTAACTGAGTTTGGCCCAATGCCAAGTAGGTAAATGA	853
QY	371	AAACAGACTCAGACATCTTCATCAATGTTTGACATATCTGACTGAACCTGCTTCTCAAGAAA	430
Db	854	AGACAGCACTGATGTTTTCATCAATCTGACAACTGCGCAATTTAGTGAAGTATCTTTTAAACCTAA	913
QY	431	ACAGAAACACCGGTTTTTTCACCTGGTCTTCTTGAACATCAATGAGTTTCCCATCAGGCAGC	490
Db	914	ACCACCTCAGAAAGTTTTTTCACAGGTTATCCTCTAATGATTAATTTCTCTATAGAGGAT	973

QY 491 CATTGACCAAGTGGTTGTTCAGTAATCTGAATATCCGTGGGACAGGTACCCACCACTTCT 550  
Db 974 TTTACCAAAACCAATATTTCTTACCAGGAGTATCTTCAAGGTGTTCCTCCATACT 1033  
QY 551 GTCTCCGCGACCGGTACGTGTTTCTGCGGACGTGGCGGAGTCAGGTGTACATGTCCTCCA 610  
Db 1034 GCAGTGGTGGTTATATATATGTCAGAGATTTGGTCCCAAGGATCTATGAAATGATGG 1093  
QY 611 AGAGCGTCCCATATCAATTAATGGAAGACGTTGTTGTGGGCTGTCCTCGAAAGGCTGA 670  
Db 1094 GTCACGTAAAACCACTCAAGTTTCAAGATGTTTATGTCGGGATCTGTTTGAATTTATTAA 1153  
QY 671 ACATCAGATGGAGGAGCTCCACTCCAGCGGACCTTTTTCAGGGGCGCTTACGGTCTT 730  
Db 1154 AAGTGACATTCATATTCAGAGACACAAATCTTTTCTTATATAGAATCCATTTGG 1213  
QY 731 CCATATGCCCTTCAGGAGGATCGTGGCTGCCACTTTCATCAAGGCTCGGACTCTCTTGG 790  
Db 1214 ATGCTGTCAACTGAGACGTGTGATTCGAGCCCATGGCTTTTCTCCRAAGGAGATCATCA 1273  
QY 791 ACTACTGGCAGG 802  
Db 1274 CTTTGGCAGG 1285

RESULT 12  
US-09-997-653-208  
; Sequence 208, Application US/09997653  
; Publication No. US20030008237A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; TITLE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: P2730PIC38  
; CURRENT APPLICATION NUMBER: US/09/997,653  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088876  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089105  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089440  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089512  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089514  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089532  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089598  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089599

; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089600  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089653  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089907  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089908  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089947  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089948  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089952  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/090246  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090252  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090254  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090349  
; PRIOR FILING DATE: 1998-06-23  
; PRIOR APPLICATION NUMBER: 60/090355  
; PRIOR FILING DATE: 1998-06-23  
; PRIOR APPLICATION NUMBER: 60/090429  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090431  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090435  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090444  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090445  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090472  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090535  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090540  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090542  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090676  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090678  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090690  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090694  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090695  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090696  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090862  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/090863  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/091360  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091478  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091544  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091519  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091626  
; PRIOR FILING DATE: 1998-07-02

; PRIOR APPLICATION NUMBER: 60/091633  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091978  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/091982  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/092182  
; PRIOR FILING DATE: 1998-07-09

Query Match 15.5%; Score 130.4; DB 9; Length 2095;  
Best Local Similarity 48.9%; Pred. No. 2e-33;  
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;

QY 20 TTITACAAAGAGAGCGGAACCTTCTTAAGCTCCAGATACAGACTGCAGCGACACCTC 79  
DB 494 TTITACAGACAAGACTTTTCACCTTCAGAGACATTCAAACCTGCTCTCATCAAAATC 553  
QY 80 CCTTCCTGCTGCTGCTGCTGACCTCATCCCAACAAAGTTGGCTGAGCGCATGGCCATCC 139  
DB 554 CATTTCTGCTGCTGCTGCTGACCTCCCAACCTTCAGATGTGAAGCCAGCGAGCCATTA 613  
QY 140 GGCAGAGCTGGGGGAAAGAGAGAGGATGGTGAAGGAAAGCAGCTGAAGACATCTTCTCTCC 199  
DB 614 GAGTTACTTGGGTGAAAAAAGTCTTGGTGGGATATGAGGTTCTTACATTTTCTTAT 673  
QY 200 TGGGGACCAACAGCAGTGCAGCGGAAGAAAGAGG-----TGGACCGAGGAGCC 250  
DB 674 TAGGCCAAAGAGGCTGAAAGAGGAAAGACAAAATGTTGGCATTTGCTTAGAGGATGAACACC 733  
QY 251 AGCGACACGGGGACATTTATCCAGAAAGATTTTCTAGAGCTCTATTACATCTGACCCCTGA 310  
DB 734 TTCTTTATGCTGACATAATCCGACAGATTTTATAGACACATATAATAACCTGACCTTGA 793  
QY 311 AGACCATGATGGGCATAGATGGTCCATCGCTTTTGTCTCAGCGGGGCTTTGTCTATGA 370  
DB 794 AAACCATTTATGGCATTCAGGTGGGTAACTGAGTTTGGCCCAATGCCAAGTACGTAATGA 853  
QY 371 AAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTCAACTGCTTCTGAAGAAA 430  
DB 854 AGACAGACACTGATGTTTTCATCAATACTGCAATTTAGTGAAGTATCTTTTAAACCTAA 913  
QY 431 ACAGACAACACAGGTTTTTCACTGGCTTCTTGAACACTCAATGAGTTTCCCATCAGCAGC 490  
DB 914 ACCACTCAGAGAAGTTTTTTCACAGGTTATCCTCTAATTGATAATTATTCTCTATAGAGAT 973  
QY 491 CATTCAGCAAGTGGTTTGTGCTAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATCT 550  
DB 974 TTTACCAAAAACCCCATATTTCTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATCT 1033  
QY 551 GCTCCGSCACCGCTACGTGTTTTCTGGCGACGTGGCGAGTCAGGTGTACAATGTCTCCA 610  
DB 1034 GCAGTGGTTGGGTTATATATATGTCAGAGATTTGGTGCCAGGATCTATGAAATGATGG 1093  
QY 611 AGAGCGTCCCATACATTAATAACTGGAAGACGTGTTTGTGGGCTCTGCCTCGAAAGCTGA 670  
DB 1094 GTCACGTAAACCCATCAAGTTTGAAGATGTTTATGTGCGGATCTGTTGATTTATTAA 1153  
QY 671 ACATCAGATTGGAGGAGCTCCACTCCAGCGACCTTTTTTCCAGGGGGCTTACGGTTCT 730  
DB 1154 AGTGAACATTCATATTCAGAGACACAAAATCTTTCTTCTATATAGAATCCATTTGG 1213  
QY 731 CCGTATGCTCTTCAGGAGGATCGTGGCTGCCACTTCATCAAGCCTCGGACTCTCTTGG 790  
DB 1214 ATGCTGTCAACTGAGACGTTGATTCAGCCCATGGCTTTTCTTCCAGGAGATCATCA 1273  
QY 791 ACTACTGGCAGG 802  
DB 1274 CTTTTTGGCAGG 1285

RESULT 13  
US-10-174-590-161  
; Sequence 161, Application US/10174590

Publication No. US20030008352A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C42  
CURRENT APPLICATION NUMBER: US/10/174,590  
CURRENT FILING DATE: 2002-06-18  
Prior application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 161  
LENGTH: 2095  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-174-590-161

Query Match 15.5%; Score 130.4; DB 9; Length 2095;  
Best Local Similarity 48.9%; Pred. No. 2e-33;  
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;

QY 20 TTATCAAGAAAGAGCGGAACTCTTAAAGCTCCAGATACAGACTCAGGACGAGACACCTC 79  
DB 494 TTTACAGACAGACTTTCACCTTCACACTTCGAGACATTCAACTGCTCTCATCAAAATC 553  
QY 80 CCTTCCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 139  
DB 554 CATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613  
QY 140 GGCAGAGCTGGGGAAG 199  
DB 614 GAGTTACTTGGGGTGAAGAAAGTCTTGGGGGATATGAGGTCTTACATTTTCTTAT 673  
QY 200 TGGGACCAACAGAGCTCAGCGGAAACGAAAGAGG-----TGGACAGGAGAGCC 250  
DB 674 TAGGCCAAGAGGCTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 733  
QY 251 AGCCAGACAGGAGGATTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 310  
DB 734 TTCTTTATGGTGACATAATCCGACAGATTTTGTAGACACATATAATAACCTGACCTTGA 793  
QY 311 AGACCATGATGGGATAG 370  
DB 794 AAACCATATGGCATTCAGGTGGGTAAGTGGTGGGATATGAGGTCTTACATTTTCTTAT 853  
QY 371 AAACAGACTCAGACATGTTTCACTCAATGTTGAGATGTTTATGTCGGGATCTGTTGAA 430  
DB 1094 GTCAGTAAACCCATCAAGTTGAGATGTTTATGTCGGGATCTGTTGAAATTTATTA 1153

QY 671 ACATCAGATTGGAGAGCTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCT 730  
DB 1154 AAGTGAACATTCATATTCAGAGAGACAAAATCTTTTCTTCTATATAGATCATTTGG 1213  
QY 731 CGGTATGCTCTTTCAGGAGGATCGTGCGCTGCCACITTCATCAAGCCCTCGGACTCTTGG 790  
DB 1214 ATGCTGTCACTGAGAGCTGTGATTGCAGCCCATGGCTTTCTTCCAAAGGAGATCATCA 1273  
QY 791 ACTACTGGCAGG 802  
DB 1274 CTTTTGGCAGG 1285

RESULT 14  
US-10-176-758-161  
Sequence 161, Application US/10176758  
Publication No. US20030008353A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C104  
CURRENT APPLICATION NUMBER: US/10/176,758  
CURRENT FILING DATE: 2002-06-21  
Prior application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 161  
LENGTH: 2095  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-176-758-161

Query Match 15.5%; Score 130.4; DB 9; Length 2095;  
Best Local Similarity 48.9%; Pred. No. 2e-33;  
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;

QY 20 TTATCAAGAAAGAGCGGAACTCTTAAAGCTCCAGATACAGACTCAGGACGAGACACCTC 79  
DB 494 TTTACAGACAGACTTTCACCTTCACACTTCGAGAGCATTCAACTGCTCTCATCAAAATC 553  
QY 80 CCTTCCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 139  
DB 554 CATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613  
QY 140 GGCAGAGCTGGGGAAG 199  
DB 614 GAGTTACTTGGGGTGAAGAAAGTCTTGGGGGATATGAGGTCTTACATTTTCTTAT 673  
QY 200 TGGGACCAACAGAGCTCAGCGGAAACGAAAGAGG-----TGGACAGGAGAGCC 250  
DB 674 TAGGCCAAGAGGCTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 733  
QY 251 AGCCAGACAGGAGGATTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 310  
DB 734 TTCTTTATGGTGACATAATCCGACAGATTTTGTAGACACATATAATAACCTGACCTTGA 793  
QY 311 AGACCATGATGGGATAG 370  
DB 794 AAACCATATGGCATTCAGGTGGGTAAGTGGTGGGATATGAGGTCTTACATTTTCTTAT 853  
QY 371 AAACAGACTCAGACATGTTTCACTCAATGTTGAGATGTTTATGTCGGGATCTGTTGAA 430

Db 854 AGACAGACACTGATGTTTTCATCACTACTGCGAATTTAGTGAAGTATCTTTTAAACCTAA 913  
QY 431 ACAGAACACACAGGTTTTTCTACTGGCTTCTGAAACTCAATGAGTTTCCCATCAGGCAGC 490  
Db 914 ACCACTCAGAGAAGTTTTTTCACAGGTTATCCCTCTAATTCATATTAATTCCTATAGAGAT 973  
QY 491 CATTCAGCAAGTGTTCAGTAATCTGAAATCCGTTGGGACAGGTACCCACCATCT 550  
Db 974 TTTACCAAAAACCATATTTCTTACCAGGATATCCTTTCAAGGTGTTCCCTCCATCT 1033  
QY 551 GCTCCGGCACCGGTACGTGTTTCTGGGACGTGGCGAGTCAAGGTGTACAATGTCCTCA 610  
Db 1034 GCAGTGGTGGGTATATATGTCAGAGATTTGGTGCAAGGATCTATGAATGATGG 1093  
QY 611 AGACGGTCCCATACATTAACACTGGAAGACGTGTTTGTGGGGCTCTCCCTCGAAAGGCTGA 670  
Db 1094 GTCACGTAACCAACCATCAAGTTTGAAGATGTTATTCGCGGATCTGTTCGAATTTATTAA 1153  
QY 671 ACATCAGATTGGAGAGCTCCACTCCCGCCGACCTTTTTTCCAGGGGCTTACGCTTCT 730  
Db 1154 AAGTGAACATTCATNTCCAGAGACACAAATCTTTTCTTCTATATAGAACTCAATTTGG 1213  
QY 731 CCGTATGCTCTTCAGGAGGATCGTGGCTGCCACTTCATCAAGCCTCGGACTCTCTGG 790  
Db 1214 ATGCTGTCACTGAGAGGTGATTCAGCCCATGGCTTTTCTTCCAGGAGATCATCA 1273  
QY 791 ACTACTGGCAGG 802  
Db 1274 CTTTTTGGCAGG 1285

RESULT 15  
US-10-063-616-35  
; Sequence 35, Application US/10063616  
; Publication No. US20030013855A1  
; GENERAL INFORMATION:  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Filvarcoff, Ellen  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Watanabe, Collin K.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3230R1C1  
; CURRENT APPLICATION NUMBER: US/10/063,616  
; CURRENT FILING DATE: 2002-05-03  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 170  
; SEQ ID NO 35  
; LENGTH: 2095  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-063-616-35

Query Match 15.5%; Score 130.4; DB 9; Length 2095;  
Best Local Similarity 48.9%; Pred. No. 2e-33;  
Matches 387; Conservative 0; Mismatches 396; Indels 9; Gaps 1;  
QY 20 TTTACAGAAGACGGGAACTTCCCTTAAGCTCCAGATACAGACTCGCAGGCAGACACCTC 79  
Db 494 TTTACAGAAGACTTTTCACATTCACACTTCGAGAGCATCAAACTGCTCTCATCAAAATC 553  
QY 80 CTTCTCCTGCTCTGCTGACCTCATCCACAAACAGTTGGCTGAGCGCATGGCCATCC 139  
Db 554 CATTTCTGGTCATTTCTGTTGACCTCCACCTTCAGATGTGAAGCCAGCGAGCCATTA 613  
QY 140 GGCAGAGCTGGGGAAGAGAGGATGGTGAAGGAAAGAGCTGAAGACATCTTCTCTCC 199  
Db 614 GAGTTACTGGGGTGAAAAAAGTCTTGGTGGGATATGAGGTTCTTACATTTTCTTAT 673

QY 200 TGGGGACCAACACAGCTGCGCGGAAACGAAAGAGG-----TGGACCAAGGAGAGCC 250  
Db 674 TAGGCCAAGAGGCTGAAAGAGAAACAAAATGTTGGCATTTGCTCTAGAGGATGAACACC 733  
QY 251 ACGGACACGGGGACATATCCCAAGGATTTCCCTAGAGCTCTATTACAACTGACCCCTGA 310  
Db 734 TTTCTTTATGGTGACATAATCCGACAGATTTTGTAGACACATATAATAACCTGACCTTGA 793  
QY 311 AGACCATGATGGGCATAGAAATGGGTGCTCATCGCTTTTGTCCCTCAGCGCGGTTTGTGATCA 370  
Db 794 AAACCATTAATGGCATTCAGGTGGGTAACTAGTTTGGCCCAATGCGCAAGTACGTAATGA 853  
QY 371 AAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTTCTGAAGAAA 430  
Db 854 AGACAGACACTGATGTTTTCATCAATCTGCAATTTAGTCAAGTATCTTTTAAACCTAA 913  
QY 431 ACAGAACACACCGGTTTTTTCACCTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGCAGC 490  
Db 914 ACCACTCAGAGAAGTTTTTTCACAGGTTATCCTCTAATTTGATAATTTATTCCTATAGAGAT 973  
QY 491 CATTCAGCAAGTGGTTTGTCAAGTAAATCTGAAATATCCGTGGGACAGGTACCCACCATCT 550  
Db 974 TTTACCAAAAACCCCATATTTCTTACCAGGATATCCTTTCAAGGTGTTCCCTCCATCT 1033  
QY 551 GCTCCGGCACCGGCTACGTGTTTCTGGCGACGTGGCGAGTCAAGGTGTACAACTGCTCCA 610  
Db 1034 GCAGTGGTGGGTATATAATGTCCAGAGATTTGGTGCCCAAGGATCTATGAATGATGG 1093  
QY 611 AGACGGTCCCATACATTAACACTGGAAGACGTGTTTGTGGGGCTCTGCTCGAAAGGCTGA 670  
Db 1094 GTCACGTAACCAACCATCAAGTTTGAAGATGTTTATGTCGGGATCTGTTTGAATTTATTAA 1153  
QY 671 ACATCAGATTGGAGAGCTCCACTCCCGCGGACCTTTTTTCCAGGGGCTTACGCTTCT 730  
Db 1154 AAGTGAACATTCATATTCAGAGACACAAATCTTTTCTTATATAGAAATCCCATTTGG 1213  
QY 731 CCGTATGCTCTTTCAGGAGGATCGTGGCTGCGCACTTCATCAAGCCTCGGACTCTCTGG 790  
Db 1214 ATGCTGTCACTGAGAGCTGATTCGACCCCATGGCTTTTCTTCCAGGAGATCATCA 1273  
QY 791 ACTACTGGCAGG 802  
Db 1274 CTTTTTGGCAGG 1285

Search completed: April 12, 2003, 03:33:46  
Job time : 66.7426 secs

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 22:52:34 ; Search time 5020.03 Seconds  
(without alignments)  
16087.603 Million cell updates/sec

Title: US-09-914-152-2

Perfect score: 2775

Sequence: 1 gtaattctcttctctgc.....aaaaaaaaaaaaaaaaaaaaa 2775

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_htg: \*  
3: gb\_in: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vl: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_om: \*  
21: em\_or: \*  
22: em\_ov: \*  
23: em\_pat: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_sts: \*  
28: em\_un: \*  
29: em\_vl: \*  
30: em\_htg\_hum: \*  
31: em\_htg\_inv: \*  
32: em\_htg\_other: \*  
33: em\_htg\_mus: \*  
34: em\_htg\_pln: \*  
35: em\_htg\_rtd: \*  
36: em\_htg\_mam: \*  
37: em\_htg\_vrt: \*  
38: em\_sy: \*  
39: em\_htgo\_hum: \*  
40: em\_htgo\_mus: \*  
41: em\_htgo\_other: \*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2775	100.0	2775	6	E38419	E38419 Novel polyp
2	2730	98.4	2762	9	AB020337	AB020337 Homo sapi
3	2330	84.0	170121	9	AF064860	AF064860 Homo sapi
4	2330	84.0	340000	9	HS21C080	AL163280 Homo sapi
5	2329	83.9	10562	6	E38420	AB041416 Homo sapi
6	933	33.6	933	9	AF145784	AF145784 Homo sapi
7	800	28.8	1576	9	AB041416	AB041416 Homo sapi
8	775	27.9	1565	9	AB041413	AB041413 Homo sapi
9	704	25.4	933	9	HS46078	AB041413 Homo sapi
10	321	11.6	1570	9	AB041415	AB041415 Pan panis
11	252	9.1	1566	9	AB041414	AB041414 Pan trogl
12	230	8.3	1360	9	AB041412	AB041412 Gorilla g
13	147	5.3	1579	9	AB041417	AB041417 Pongo pyg
14	53	1.9	1746	9	HS800208	AL049423 Homo sapi
15	53	1.9	2771	9	BC013407	BC013407 Homo sapi
16	53	1.9	2828	9	AF217982	AF217982 Homo sapi
17	51	1.8	3114	9	BC032365	BC032365 Homo sapi
18	51	1.8	155847	9	AC092766	AC092766 Homo sapi
19	50	1.8	647	9	AK024747	AK024747 Homo sapi
20	50	1.8	647	9	AK026570	AK026570 Homo sapi
21	50	1.8	899	9	BC014163	BC014163 Homo sapi
22	50	1.8	972	10	BC030079	BC030079 Mus muscu
23	50	1.8	1032	4	AF013214	AF013214 Bos tauru
24	50	1.8	1069	6	AR177949	AR177949 Sequence
25	50	1.8	1069	8	CAR400862	AJ400862 Cicer ari
26	50	1.8	1132	10	BC011291	BC011291 Mus muscu
27	50	1.8	1138	10	BC011106	BC011106 Mus muscu
28	50	1.8	1391	10	BC016602	BC016602 Mus muscu
29	50	1.8	1474	6	A65341	A65341 Sequence 64
30	50	1.8	1474	6	AR150479	AR150479 Sequence
31	50	1.8	1480	9	BC032007	BC032007 Homo sapi
32	50	1.8	1586	6	A76337	A76337 Sequence 3
33	50	1.8	1586	8	BNPL3	X73850 B.napus (pN
34	50	1.8	1794	10	BC029222	BC029222 Mus muscu
35	50	1.8	2022	9	AK000266	AK000266 Homo sapi
36	50	1.8	2046	9	BC016024	BC016024 Homo sapi
37	50	1.8	2050	9	BC002967	BC002967 Homo sapi
38	50	1.8	2301	6	AR168877	AR168877 Sequence
39	50	1.8	2301	6	AR193664	AR193664 Sequence
40	50	1.8	2301	6	AX101316	AX101316 Sequence
41	50	1.8	2301	6	AK074450	AK074450 Homo sapi
42	50	1.8	2374	5	XLA5870	AJ005870 Xenopus l
43	50	1.8	2710	6	AR168892	AR168892 Sequence
44	50	1.8	2710	6	AR168905	AR168905 Sequence
45	50	1.8	2710	6	AR193679	AR193679 Sequence

ALIGNMENTS

RESULT 1

E38419  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

E38419  
Novel polypeptide.  
E38419  
E38419.1 GI:18626993  
JP 2000245464-A/1.  
Homo sapiens.  
Homo sapiens

2775 bp DNA linear PAT 31-JAN-2002

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2775)  
Narimatsu,H., Isshiki,S., Togayauchi,A. and Sasaki,K.  
Novel polypeptide  
Patent: JP 2000245464-A 1 12-SEP-2000;

Pred. No. is the number of results predicted by chance to have a



Db	1741	CCTGACTTGAACCACTCCACAGTGTGCTCCCTTAGGAGGGGACACTGATGATGAGGTC	1800
Qy	1801	TCGGAGCGGCATCTTCCATCCCTGTGAGTCCCTCCACCTCAGCTCCAGTCCCTTGT	1860
Db	1801	TCGGAGCGGCATCTTCCATCCCTGTGAGTCCCTCCACCTCAGCTCCAGTCCCTTGT	1860
Qy	1861	GCTTTTGGAGCTAAGCTGGGATGACCAAAATCACCCAGGTCCTTCAATTCACAGGCT	1920
Db	1861	GCTTTTGGAGCTAAGCTGGGATGACCAAAATCACCCAGGTCCTTCAATTCACAGGCT	1920
Qy	1921	GGATCTAGCTGGGATTGAGTCCATGTTATCGCTCGGTACTCAACACACACCCCAAGTTTCA	1980
Db	1921	GGATCTAGCTGGGATTGAGTCCATGTTATCGCTCGGTACTCAACACACACCCCAAGTTTCA	1980
Qy	1981	TCCGAGGAAATCCCCGAGTGGATGACAGCTCACAATGCTGAGGAACACCCAGCTCTGGA	2040
Db	1981	TCCGAGGAAATCCCCGAGTGGATGACAGCTCACAATGCTGAGGAACACCCAGCTCTGGA	2040
Qy	2041	CAGAGTTCTTATAATGTATAAATAGGCTCAGAAACCACTGCATTCTGACCTGCTGTAC	2100
Db	2041	CAGAGTTCTTATAAATGTATAAATAGGCTCAGAAACCACTGCATTCTGACCTGCTGTAC	2100
Qy	2101	AGACTGCCACACTGCTGCACCTGCCATAGGAGCAGGACATCCCTTCTGAGCCATCTGCTG	2160
Db	2101	AGACTGCCACACTGCTGCACCTGCCATAGGAGCAGGACATCCCTTCTGAGCCATCTGCTG	2160
Qy	2161	CTCTCTCATTTCAATCACCCCAACTGTCCCTTGTGATCAATGGGGACAGCCACTGC	2220
Db	2161	CTCTCTCATTTCAATCACCCCAACTGTCCCTTGTGATCAATGGGGACAGCCACTGC	2220
Qy	2221	CCCAGGACACTTTAGGCTCTCAGTTCAACTGAAGGACAGTTGAACTCAGATGGGTT	2280
Db	2221	CCCAGGACACTTTAGGCTCTCAGTTCAACTGAAGGACAGTTGAACTCAGATGGGTT	2280
Qy	2281	CATGTGGGATCTGGGAGCTTCTGGGAATTCAGTTGGAGTCAAGTCAGGATGCTCTCAA	2340
Db	2281	CATGTGGGATCTGGGAGCTTCTGGGAATTCAGTTGGAGTCAAGTCAGGATGCTCTCAA	2340
Qy	2341	GGACCCCTCGGGCTCAGAGCCCTAAAGTGGGCCCTGGTGAAGCAGGGTGGTCTCGGTC	2400
Db	2341	GGACCCCTCGGGCTCAGAGCCCTAAAGTGGGCCCTGGTGAAGCAGGGTGGTCTCGGTC	2400
Qy	2401	ACTTCCCAAGCTGAGCCCAAGCTCATCTTGAATGTCATTTGGCCGAGGACAAC	2460
Db	2401	ACTTCCCAAGCTGAGCCCAAGCTCATCTTGAATGTCATTTGGCCGAGGACAAC	2460
Qy	2461	TGAACCTTGTGTTGCTGTTTAGCCTTCAGTTTCTCGCTGCCTCCTACCCAGAGTT	2520
Db	2461	TGAACCTTGTGTTGCTGTTTAGCCTTCAGTTTCTCGCTGCCTCCTACCCAGAGTT	2520
Qy	2521	TGTCGGACCTCTGTTGCAGGGTTCGTATAAACCAAGTACTTCGTTAGTTTGGCCATT	2580
Db	2521	TGTCGGACCTCTGTTGCAGGGTTCGTATAAACCAAGTACTTCGTTAGTTTGGCCATT	2580
Qy	2581	CAGCCATGGTCACGTGACATGCAAAAGTAACTCTGTCTCTAAATATAGAAATGATTTTCT	2640
Db	2581	CAGCCATGGTCACGTGACATGCAAAAGTAACTCTGTCTCTAAATATAGAAATGATTTTCT	2640
Qy	2641	TTTAATTTTTTACTTTACAGACTTTACTTTGTACTCAGAGAGGCGCTCATGCGTG	2700
Db	2641	TTTAATTTTTTACTTTACAGACTTTACTTTGTACTCAGAGAGGCGCTCATGCGTG	2700
Qy	2701	TGTCACATATAAATGTTGCACTTAACTCTTAAAAAATAAAAAAAAAAAAAAAAAAAAA	2760
Db	2701	TGTCACATATAAATGTTGCACTTAACTCTTAAAAAATAAAAAAAAAAAAAAAAAAAAA	2760
Qy	2761	AAAAAAAAAAAAAAAA 2775	
Db	2761	AAAAAAAAAAAAAAAA 2775	

RESULT 2  
AB020337

LOCUS	AB020337	2762 bp	mRNA	linear	PRI 14-APR-2000
DEFINITION	Homo sapiens mRNA for UDP-Gal:GlcNAc betal,3-galactosyltransferase 5, complete cds.				
ACCESSION	AB020337				
VERSION	AB020337.1	GI:4835502			
KEYWORDS	UDP-Gal:GlcNAc betal,3-galactosyltransferase 5.				
SOURCE	Homo sapiens Adenocarcinoma cell_line:Colo 205 cDNA to mRNA.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Isshiki,S., Togayachi,A., Kudo,T., Nishihara,S., Watanabe,M., Kubota,T., Kitajima,M., Shiraishi,N., Sasaki,K., Andoh,T. and Narimatsu,H.				
TITLE	Cloning, expression, and characterization of a novel beta1,3-galactose:beta-N-acetylglucosamine beta1,3-galactosyltransferase (beta3Gal-T5) responsible for synthesis of type I chain in colorectal and pancreatic epithelia and tumor cells derived therefrom				
JOURNAL	J. Biol. Chem. 274 (18), 12499-12507 (1999)				
MEDLINE	99230269				
REFERENCE	2 (bases 1 to 2762)				
AUTHORS	Isshiki,S., Togayachi,A. and Narimatsu,H.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-Nov-1998) Hisashi Narimatsu, Soka University, Institute of Life Science, 1-236, Tangi-cho, Hachioji, Tokyo 192-8577, Japan (E-mail:sissiki@po.llnwt.or.jp, Tel:81-426-91-9466, Fax:81-426-91-9315)				
FEATURES	Location/Qualifiers				
source	1..2762				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/chromosome="21"				
	/map="21q22.3"				
	/cell_line="Colo 205"				
	/cell_type="Adenocarcinoma"				
gene	1..2762				
	/gene="betal,3-Galt 5"				
exon	1..273				
	/number=1				
exon	274..433				
	/gene="betal,3-Galt 5"				
	/number=3				
exon	434..2762				
	/gene="betal,3-Galt 5"				
	/number=4				
CDS	434..1366				
	/gene="betal,3-Galt 5"				
	/codon_start=1				
	/product="UDP-Gal:GlcNAc betal,3-galactosyltransferase 5"				
	/protein_id="BAA77664.1"				
	/db_xref="GI:4835503"				
	/translation="MAFFKMRLMYICLLVLGLALCYFSMSYSLNPFKQSFVKKDGNF LKLPTDCRQTPEFLIVLTSSHKQLAERAIROTGWKRMVKQKLIKTLFTLGTTS AAETKEVDQSORHGDIQKDELFDVYNLTAKTMGIEMWHRFCPOAFVNMKTSDMF INVDYELLKLNRTTREFDFKLNEFPRIQPPFSKMVSXSVYPMDRYPPFGSGTG YVESGDVASOVNVSKSPVYIKLEDVFVGLCLRLNLRLBELHSQPIFFPGGLRFSCV LFRRIVACHFIKPRILLDWQALESNGEDCPVP"				
	2762				
polyA_site	/gene="betal,3-Galt 5"				
	/note="45 a nucleotides"				
BASE COUNT	651 a 703 c 679 g 729 t				
ORIGIN					
Query Match	98.4%	Score 2730;	DB 9;	Length 2762;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 2730;	Conservative	0;	Mismatches	0;	Gaps 0;
Qy	1 GTGAATTCCTCTTTCTCTCGTAGCTGGGATATTCTTCTCTCCTGCCCTGGACATCA	60			
Dd	33 GTGAATTCCTCTTTCTCTCGTAGCTGGGATATTCTTCTCTCCTGCCCTGGACATCA	92			





Db 2253 CCCAGGACACTTTAGGGCTCTCAGTTCAACTGAAGGACAGTGAAGTCAAGTGGGTT 2312

Qy 2281 CATGTGGATTCTGGGAGCTTTCTGGGAATTCAGTTGGAGTCAAGTCAAGTGGTCTCAA 2340

Db 2313 CATGTGGATTCTGGGAGCTTTCTGGGAATTCAGTTGGAGTCAAGTCAAGTGGTCTCAA 2372

Qy 2341 GGACCCCTCGGCTCAGAGCCCTAAAGTGGGCTGGTGAAGCAGGAGTGGTCTGGCTCC 2400

Db 2373 GGACCCCTCGGCTCAGAGCCCTAAAGTGGGCTGGTGAAGCAGGAGTGGTCTGGCTCC 2432

Qy 2401 ACTTCCCAAGCTGAGCCAAAGCTCATCTTCATTGAATGCTCTATTGGCCGAGGAACAAC 2460

Db 2433 ACTTCCCAAGCTGAGCCAAAGCTCATCTTCATTGAATGCTCTATTGGCCGAGGAACAAC 2492

Qy 2461 TGAATTTGTGGTGTGCTTTAGCCTTCAGTTTGGCTCCGCTCCCTACCCAGAGGTT 2520

Db 2493 TGAATTTGTGGTGTGCTTTAGCCTTCAGTTTGGCTCCGCTCCCTACCCAGAGGTT 2552

Qy 2521 TGTGGAGCCCTGTGTGAGGCTTGTATAAACCAAGGACTCTCGTAGTTTGGCCCAT 2580

Db 2553 TGTGGAGCCCTGTGTGAGGCTTGTATAAACCAAGGACTCTCGTAGTTTGGCCCAT 2612

Qy 2581 CAGCCATGCTCAGCGAGATCAAGTAATCTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2640

Db 2613 CAGCCATGCTCAGCGAGATCAAGTAATCTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2672

Qy 2641 TTTAATTTTACTTTACGAGCTTTACTTTGCTTACTCAGAGAGAGGCTTCACATGGCTG 2700

Db 2673 TTTAATTTTACTTTACGAGCTTTACTTTGCTTACTCAGAGAGAGGCTTCACATGGCTG 2732

Qy 2701 TGTCAATATAAAGTTGGACTAAACTCTT 2730

Db 2733 TGTCAATATAAAGTTGGACTAAACTCTT 2762

RESULT 3

AF064860 170121 bp DNA linear PRI 05-MAR-2002

LOCUS Homo sapiens chromosome 21 clone PAC 70124 map 21q22.3, complete

DEFINITION sequence.

ACCESSION AF064860

VERSION AF064860.2 GI:18958624

KEYWORDS HTG; HTGS-DRAFT.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T., Park,H.-S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.-K., Soeda,E., Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K., Poller,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R., Patterson,D., Reichwald,K., Rump,A., Schillhabel,M.B., Schudy,A., Zimmermann,W., Rosenthal,A., Kudoh,J., Kawasaki,K., Asakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuyama,S., Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G., Hornischer,K., Brandt,P., Scharfe,M., Schoen,O., Desario,A., Hennig,S., Riesselmann,L., Dagand,E., Haef,T., Wehrmeyer,S., Borzym,K., Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and Yaspo,M.Laure.

The DNA sequence of human chromosome 21

Nature 405 (6784), 311-319 (2000)

20289799

10830953

REFERENCE 2 (bases 1 to 170121)

Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B., Weber,J., Schattevoy,R., Yaspo,M.-L. and Rosenthal,A.

DIRECT SUBMISSION

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

Submitted (12-MAY-1998) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany

3 (bases 1 to 170121)

Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B.,

Weber,J., Schattevoy,R., Yaspo,M.-L., Rosenthal,A., Yaspo,M.-L. and Rosenthal,A.

DIRECT SUBMISSION

TITLE

JOURNAL

Submitted (27-FEB-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany

4 (bases 1 to 170121)

Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B., Weber,J., Schattevoy,R., Yaspo,M.-L. and Rosenthal,A.

DIRECT SUBMISSION

TITLE

JOURNAL

Submitted (05-MAR-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany

On Feb 27, 2002 this sequence version replaced gi:3171153.

Location/Qualifiers

1. .170121

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="21"

/map="21q22.3"

/clone="PAC 70124"

BASE COUNT 45184 a 36756 c 37509 g 50672 t

ORIGIN

Query Match 84.0%; Score 2330; DB 9; Length 170121;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 402 ATGGCTTTCCCGAGATGACATGATGATATATTTGCTTCTGGTCTCGGGCTCTTCT 461

Db 90048 ATGGCTTTCCCGAGATGACATGATGATATATTTGCTTCTGGTCTCGGGCTCTTCT 90107

Qy 462 TTGTATTTAGCATGTACAGTCTAAATCCTTTCAAAGAACAGTCTTTGTTTACAAGAAA 521

Db 90108 TTGTATTTAGCATGTACAGTCTAAATCCTTTCAAAGAACAGTCTTTGTTTACAAGAAA 90167

Qy 522 GACGGAACTTCCTTAAGCTCCCGAGATACAGTCTGAGGACAGACCTCCCTTCTCGTCT 581

Db 90168 GACGGAACTTCCTTAAGCTCCCGAGATACAGTCTGAGGACAGACCTCCCTTCTCGTCT 90227

Qy 582 CTGCTGGTGAACCTCATCCCAACAGTGTGGTGAAGGCGATGGCCATCCGGCAGAGTGG 641

Db 90228 CTGCTGGTGAACCTCATCCCAACAGTGTGGTGAAGGCGATGGCCATCCGGCAGAGTGG 90287

Qy 642 GGGAAAGAGAGGATGTTGAGGGAAGAGCAGTGAAGACATTTCTTCTCTCTGGGACCACC 701

Db 90288 GGGAAAGAGAGGATGTTGAGGGAAGAGCAGTGAAGACATTTCTTCTCTCTGGGACCACC 90347

Qy 702 AGCAGTGCAGCGGAAAGAGAGTGGACAGGAGGCGACAGACAGCGGACATTTATTC 761

Db 90348 AGCAGTGCAGCGGAAAGAGAGTGGACAGGAGGCGACAGACAGCGGACATTTATTC 90407

Qy 762 CAGAAGGATTTCTTAGACGCTTATTACAATCTGACCTGGAAGACCATGATGGGCATAGAA 821

Db 90408 CAGAAGGATTTCTTAGACGCTTATTACAATCTGACCTGGAAGACCATGATGGGCATAGAA 90467

Qy 822 TGGGTCATCGCTTTTGTCTCAGGGGGCGCTTTGTGATGAAGAAACAGACTCAGACATGTC 881

Db 90468 TGGGTCATCGCTTTTGTCTCAGGGGGCGCTTTGTGATGAAGAAACAGACTCAGACATGTC 90527

Qy 882 ATCAATGTGTACTATCTGACTGAACCTGCTTCTGAAGAAAACAGAACACAGGTTTTC 941

Db 90528 ATCAATGTGTACTATCTGACTGAACCTGCTTCTGAAGAAAACAGAACACAGGTTTTC 90587

Qy 942 ACTGGCTTCTTGAAACTCAATGAGTTTCCCATCAGGAGGACCATTCAGCAAGTGGTTTCTC 1001

Db 90588 ACTGGCTTCTTGAAACTCAATGAGTTTCCCATCAGGAGGACCATTCAGCAAGTGGTTTCTC 90647

Qy 1002 AGTAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGGACCGGCTACGCTG 1061

Db 90648 AGTAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGGACCGGCTACGCTG 90707

Qy 1062 TTTTCTGGCGAGTGGCGAGTCAAGTGTACAATGTCTCCAAAGAGCGTCCCATACATATAA 1121

Db 90708 TTTTCTGGCGAGTGGCGAGTCAAGTGTACAATGTCTCCAAAGAGCGTCCCATACATATAA 90767

QY	1122	CTGGAAGACGTGTTTGTGGGGCTCTGCCTCGAAAGGCTGAACATCAGATTGGAGGAGCTC	1181
Db	90768	CTGGAAGACGTGTTTGTGGGGCTCTGCCTCGAAAGGCTGAACATCAGATTGGAGGAGCTC	90827
QY	1182	CACCTCCAGCCGACCTTTTTCAGAGGGGCTTACGCTTCTCCGATGCTCTTCAGGAGG	1241
Db	90828	CACCTCCAGCCGACCTTTTTCAGAGGGGCTTACGCTTCTCCGATGCTCTTCAGGAGG	90887
QY	1242	ATCGTGGCTGCCACTTCATCAGCCTCGGACTCTCTTGGACTACTGCGAGGCTCTAGAG	1301
Db	90888	ATCGTGGCTGCCACTTCATCAGCCTCGGACTCTCTTGGACTACTGCGAGGCTCTAGAG	90947
QY	1302	AATTCGGGGGGAAGATTCTCGCCCTGCTGAGGGAGCCAGAGGACACATCCGAGCAA	1361
Db	90948	AATTCGGGGGGAAGATTCTCGCCCTGCTGAGGGAGCCAGAGGACACATCCGAGCAA	91007
QY	1362	GTTCAGATAAACCCGCTGGGATAGTTTTTGTAGATTTTGAAGAGGGGGGACAGAG	1421
Db	91008	GTTCAGATAAACCCGCTGGGATAGTTTTTGTAGATTTTGAAGAGGGGGGACAGAG	91067
QY	1422	GATGCTGTTCTTCAGTGCTGAAATCCAGCCAGAAATGTCGGTGTTCATGAAGTCACTGAT	1481
Db	91068	GATGCTGTTCTTCAGTGCTGAAATCCAGCCAGAAATGTCGGTGTTCATGAAGTCACTGAT	91127
QY	1482	TAGTTCCTCCACTTGGTGGCCCGCAGCAATAATAGGCCGCTCTTGGGACGACACTCTTC	1541
Db	91128	TAGTTCCTCCACTTGGTGGCCCGCAGCAATAATAGGCCGCTCTTGGGACGACACTCTTC	91187
QY	1542	ATACTAAGTGTTCAGATACACACTGGATTTTTTGCATTTTTCAGGGTTCAGTATCCCTATGACA	1601
Db	91188	ATACTAAGTGTTCAGATACACACTGGATTTTTTGCATTTTTCAGGGTTCAGTATCCCTATGACA	91247
QY	1602	TGATGGGTGTTACCATCTTAATTTTACAGGCAAGGACACAGCAGCTGGAGAGGTACAGA	1661
Db	91248	TGATGGGTGTTACCATCTTAATTTTACAGGCAAGGACACAGCAGCTGGAGAGGTACAGA	91307
QY	1662	RACTTGTCACAGGCTCACAGCCAGTAGGATAGGAGGGGAATGAATAATCGAGCACTGT	1721
Db	91308	RACTTGTCACAGGCTCACAGCCAGTAGGATAGGAGGGGAATGAATAATCGAGCACTGT	91367
QY	1722	CAGATCTGCTGGGACGCCCTGACTTGAACCACTCCACAGCTGCTGCCCTCTAGGAGG	1781
Db	91368	CAGATCTGCTGGGACGCCCTGACTTGAACCACTCCACAGCTGCTGCCCTCTAGGAGG	91427
QY	1782	GGACACTGATGATGAGGTCTCGGAGCGCGCATCCTCCATCCCTGTCGAGTCCCTCCAC	1841
Db	91428	GGACACTGATGATGAGGTCTCGGAGCGCGCATCCTCCATCCCTGTCGAGTCCCTCCAC	91487
QY	1842	CTCAGTCCCGAGTCTGTGCTTTTGGAGCTAAGCTGGGATGACCAAAATTCACCCAG	1901
Db	91488	CTCAGTCCCGAGTCTGTGCTTTTGGAGCTAAGCTGGGATGACCAAAATTCACCCAG	91547
QY	1902	CTCCTTCATTCACAGGCTGGATGTAGCTGGGATGAGTCCATCTTATCGGCTCGGTACT	1961
Db	91548	CTCCTTCATTCACAGGCTGGATGTAGCTGGGATGAGTCCATCTTATCGGCTCGGTACT	91607
QY	1962	CAACACACCCCAAGTTTCATCCGAGGAATATGTCGCCAGTGGATGAGCTCAGTCACTGCTG	2021
Db	91608	CAACACACCCCAAGTTTCATCCGAGGAATATGTCGCCAGTGGATGAGCTCAGTCACTGCTG	91667
QY	2022	AGGACACCCAGCTCGGACAGAGTCTTTATAATGATTAATTAATGAGTCTCAGAACCACT	2081
Db	91668	AGGACACCCAGCTCGGACAGAGTCTTTATAATGATTAATTAATGAGTCTCAGAACCACT	91727
QY	2082	GCATCTGACCTGCTACAGACTGCCACACTGCTGACCTGCTAGGAGCAGGACATC	2141
Db	91728	GCATCTGACCTGCTACAGACTGCCACACTGCTGACCTGCTAGGAGCAGGACATC	91787
QY	2142	CCTTCTGACCACTCTGCTCTCTCATTTTCACACCCCAACTCTCCCTCTGTTTTGATC	2201
Db	91788	CCTTCTGACCACTCTGCTCTCTCATTTTCACACCCCAACTCTCCCTCTGTTTTGATC	91847

QY	2202	AATGGGACCGACCACTGCCAGGAGCACTTTAGGGCTCTCAGTTCAAACTGAAGACA	2261
Db	91848	AATGGGACCGACCACTGCCAGGAGCACTTTAGGGCTCTCAGTTCAAACTGAAGACA	91907
QY	2262	GTTGAACCTCAGATGGGTTTCATGTGGGATTCCTGGGAGCTTCTGGGAATTCAGTTGAGT	2321
Db	91908	GTTGAACCTCAGATGGGTTTCATGTGGGATTCCTGGGAGCTTCTGGGAATTCAGTTGAGT	91967
QY	2322	CAAGTCAGGATGCTCTCAAGGACCCCTCGGGCTCAGAGCCCTAAAGTGGCCCTGGTGA	2381
Db	91968	CAAGTCAGGATGCTCTCAAGGACCCCTCGGGCTCAGAGCCCTAAAGTGGCCCTGGTGA	92027
QY	2382	GCAGGTGGTCTCGGTCCTCACTTCCAGCCCTGAGCCCAAGCTCATCTTCAATGCTCT	2441
Db	92028	GCAGGTGGTCTCGGTCCTCACTTCCAGCCCTGAGCCCAAGCTCATCTTCAATGCTCT	92087
QY	2442	CATTGGCCGAGGAACAACCTGAACCTTGTGGTTCCTGCTTTAGCCCTTCAATGCTCT	2501
Db	92088	CATTGGCCGAGGAACAACCTGAACCTTGTGGTTCCTGCTTTAGCCCTTCAATGCTCT	92147
QY	2502	TGCTCTCTACCCAGGAGTTTGTGCGAGGCTGTGTTGACAGGTTGTATAAAACCAAGTAC	2561
Db	92148	TGCTCTCTACCCAGGAGTTTGTGCGAGGCTGTGTTGACAGGTTGTATAAAACCAAGTAC	92207
QY	2562	TTCCTTATGTTTGGCCATTCAGCCATGTCAGTGCATGACATGCAAGTAATCTTCTCTAA	2621
Db	92208	TTCCTTATGTTTGGCCATTCAGCCATGTCAGTGCATGACATGCAAGTAATCTTCTCTAA	92267
QY	2622	TTATAGAAATGATTTTCTTTTAAATTTTACTTTTACCAGACTTTTACTTTGCTACTCAG	2681
Db	92268	TTATAGAAATGATTTTCTTTTAAATTTTACTTTTACCAGACTTTTACTTTGCTACTCAG	92327
QY	2682	AAGAGGCTCAGTGGCTGTGCATATATAATGTTGGACTAAACTCTTA	2731
Db	92328	AAGAGGCTCAGTGGCTGTGCATATATAATGTTGGACTAAACTCTTA	92377

RESULT 4  
HS21C080  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

HS21C080 340000 bp DNA linear PRI 24-MAY-2000  
Homo sapiens chromosome 21 segment HS21C080.  
AL163280 AP001735 BA000005  
AL163280.2 GI:7717369

human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,  
Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,  
Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,  
Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,  
Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,  
Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K.,  
Asakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuyma,S.,  
Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G.,  
Hornischer,K., Brandt,P., Scharfe,M., Schoen,O., Desario,A.,  
Reichelt,J., Kauer,G., Bloeker,H., Ramser,J., Beck,A., Klages,S.,  
Hennig,S., Riesselmann,L., Dagand,E., Wehrmeyer,S., Borzym,K.,  
Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and  
Yaspo,M.L.

Direct Submission  
Submitted (05-MAY-2000) The Chromosome 21 Mapping and Sequencing  
Consortium: \* RIKEN Genomic Sciences Center, Human Genome Research  
Group \* Institute of Molecular Biotechnology, Genome Analysis \*  
Keio University School of Medicine, Dept. of Molecular Biology \*  
GBF, Dept. of Genome Analysis \* Max-Planck Institute for Molecular  
Genetics (addresses see below)  
The Chromosome 21 Mapping and Sequencing Consortium consists of  
\* RIKEN Genomic Sciences Center, Human Genome Research Group, \*  
Sagamihara 228-8555, Japan,  
\* e.mail: sakaki@gsc.riken.go.jp

FEATURES	source	* GBF, Dept. of Genome Analysis, * Keio University School of Medicine, Dept. of Molecular Biology, * Tokyo 160-8582, Japan, * e.mail: shimizudmb-med.keio.ac.jp * URL: http://adenine.dmb.med.keio.ac.jp/ and * Keio University School of Medicine, Dept. of Molecular Biology, * Tokyo 160-8582, Japan, * e.mail: shimizudmb-med.keio.ac.jp * URL: http://adenine.dmb.med.keio.ac.jp/ and * GBF, Dept. of Genome Analysis, * Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e.mail: info.genome@gbf.de * URL: http://genome.gbf.de/ and * Max-Planck Institute for Molecular Genetics, * Ihnestrasse 73, D-14195 Berlin, Germany, * e.mail: info.chr21@olgen.mpg.de * URL: http://chr21.rz-berlin.mpg.de/ Location/Qualifiers 1. 340000 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="21" /map="21q22.3" <1. 125946 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="21" /map="21q22.3" /clone="P160L9, 5' partial" /clone.lib="RPC11.3-5 PAC library" /note="Accession No. AF121897" 73410. 243533 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="21" /map="21q22.3" /clone="P70124" /clone.lib="RPC11.3-5 PAC library" /note="Accession No. AF064860" 224137. 340000 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="21" /map="21q22.3" /clone="P70124" /clone.lib="RPC11.3-5 PAC library" /note="Accession No. AF121782" 286628. 340000 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="21" /map="21q22.3" /clone="BAC-291B3, 3' partial" /clone.lib="BAC library" /note="Accession No. AF064857" 1088. 1566 /note="L1M4" /rpt_family="LINE/L1" /rpt_type=DISPERSED 2398. 2495 /note="(TCTA)n" /rpt_family="Simple_repeat" /rpt_type=TANDEM complement(2495. 2878) /note="THEIC" /rpt_family="LTR/MaLR" /rpt_type=DISPERSED 2879. 2931 /note="(TCTA)n"	repeat_region	/rpt_family="Simple_repeat" /rpt_type=TANDEM 2932. 3004 /note="(TCCA)n" /rpt_family="Simple_repeat" /rpt_type=TANDEM 3005. 3074 /note="(TCCA)n" /rpt_family="Simple_repeat" /rpt_type=TANDEM 3473. 19157 /gene="SH3BGR" join(<3473. 3565,12649. 12678,15331. 15460,18739. 19157) /gene="SH3BGR" join(3473. 3565,12649. 12678,15331. 15426) /partial /gene="SH3BGR" /note="Accession No. X93498" /codon_start=1 /product="21-Glutamic Acid Rich protein 21-GARP" /protein_id="CAB90445.1" /db_xref="GI:7717370" /translation="GSEKAEGGETEAKGSEVDGNLPEAQEKNEEGETATETETE IAMEGAEAEAEETAEGEPEGEIDS" 3473. 3565 /gene="SH3BGR" /number=4 5542. 5829 /note="AluSx" /rpt_family="SINE/Alu" /rpt_type=DISPERSED complement(6887. 7067) /note="L1MD3" /rpt_family="LINE/L1" /rpt_type=DISPERSED complement(8482. 8762) /note="AluJo" /rpt_family="SINE/Alu" /rpt_type=DISPERSED complement(8865. 8964) /note="L2" /rpt_family="LINE/L2" /rpt_type=DISPERSED 9131. 9424 /note="AluSq" /rpt_family="SINE/Alu" /rpt_type=DISPERSED complement(9723. 9882) /note="L2" /rpt_family="LINE/L2" /rpt_type=DISPERSED complement(9989. 10077) /note="L2" /rpt_family="LINE/L2" /rpt_type=DISPERSED complement(10078. 10429) /note="THE1B" /rpt_family="LTR/MaLR" /rpt_type=DISPERSED complement(10430. 10651) /note="L2" /rpt_family="LINE/L2" /rpt_type=DISPERSED 10980. 11000 /note="(TTTA)n" /rpt_family="Simple_repeat" /rpt_type=TANDEM complement(11002. 11282) /note="AluSp" /rpt_family="SINE/Alu" /rpt_type=DISPERSED complement(11315. 11556) /note="AluJo"

repeat_region	/rpt_family="SINE/Alu"	
	/rpt_type=DISPERSED	
exon	11702..11753	
	/note="MIR"	
repeat_region	/rpt_family="SINE/MIR"	
	/rpt_type=DISPERSED	
repeat_region	12649..12678	
	/gene="SH3BGR"	
repeat_region	/number=5	
	complement(13115..13425)	
repeat_region	/note="AluSx"	
	/rpt_family="SINE/Alu"	
repeat_region	/rpt_type=DISPERSED	
	complement(13398..13892)	
repeat_region	/note="Alu"	
	/rpt_family="SINE/Alu"	
repeat_region	/rpt_type=DISPERSED	
	14013..14131	
repeat_region	/note="(CA)n"	
	/rpt_family="Simple_repeat"	
Query Match 84.0%; Score 2330; DB 9; Length 340000;		
Best Local Similarity 100.0%; Pred. No. 0;		
Matches 2330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY 402	ATGCGCTTCCGAGAGATGAGATGTATATTTGCGCTTCTGGTCTGGGGCTCTTTGT	461
DB 163456	ATGCGCTTCCGAGAGATGAGATGTATATTTGCGCTTCTGGTCTGGGGCTCTTTGT	163515
QY 462	TTGATATTTTACATGTACAGTCTAATCTTTCAGAGACAGTCCCTTGTTTACAGAAA	521
DB 163516	TTGATATTTTACATGTACAGTCTAATCTTTCAGAGACAGTCCCTTGTTTACAGAAA	163575
QY 522	GACGGGAATTCCTTAAGTCCAGATACAGACTGCGAGCAGACACCTCCCTTCCTCGTC	581
DB 163576	GACGGGAATTCCTTAAGTCCAGATACAGACTGCGAGCAGACACCTCCCTTCCTCGTC	163635
QY 582	CTGCTGGTGAACCTATCCCAAAAGAGTTGGCTGAGCGCATGCCATCCGGCAGAGTGG	641
DB 163636	CTGCTGGTGAACCTATCCCAAAAGAGTTGGCTGAGCGCATGCCATCCGGCAGAGTGG	163695
QY 642	GGGAAAGAGAGATGGTCAAGGGAAGAGCTGAAGACATTTCTCTCTGGGGACACC	701
DB 163696	GGGAAAGAGAGATGGTCAAGGGAAGAGCTGAAGACATTTCTCTCTGGGGACACC	163755
QY 702	AGCAGTGCAGGGGAAACAAAGAGGTGGACAGGAGAGCCAGGACACGGGACATTATC	761
DB 163756	AGCAGTGCAGGGGAAACAAAGAGGTGGACAGGAGAGCCAGGACACGGGACATTATC	163815
QY 762	CAGAAGGATTTCCCTAGAGCTATATTACATCTGACCCCTGAAGACCATGATGGCATAGAA	821
DB 163816	CAGAAGGATTTCCCTAGAGCTATATTACATCTGACCCCTGAAGACCATGATGGCATAGAA	163875
QY 822	TGGTCCATCGCTTTTGTCTCAGCGGGGTTTGTGATGAAACAGACTCAGACATGTTT	881
DB 163876	TGGTCCATCGCTTTTGTCTCAGCGGGGTTTGTGATGAAACAGACTCAGACATGTTT	163935
QY 882	ATCAATGTTGACTATCTGACTGAAGTCTCTGAGAAAACAGACACACCGAGTTTTC	941
DB 163936	ATCAATGTTGACTATCTGACTGAAGTCTCTGAGAAAACAGACACACCGAGTTTTC	163995
QY 942	ACTGGCTTCTTGAACACTCAATGAGTTTCCCATCAGCAGCCCATTCAGCAAGTGTGTTC	1001
DB 163996	ACTGGCTTCTTGAACACTCAATGAGTTTCCCATCAGCAGCCCATTCAGCAAGTGTGTTC	164055
QY 1002	AGTAAATCTGAATATCCGTGGGACAGTACCACCATTTCTGCTCCGGCACCAGCTACGTG	1061
DB 164056	AGTAAATCTGAATATCCGTGGGACAGTACCACCATTTCTGCTCCGGCACCAGCTACGTG	164115
QY 1062	TTTCTTGGCAGCTGGCAGTCAAGTGTACATGTCTCCAAAGAGGTCCCATACATATAA	1121
DB 164116	TTTCTTGGCAGCTGGCAGTCAAGTGTACATGTCTCCAAAGAGGTCCCATACATATAA	164175

QY 1122	CTGGAAGACGTGTTTGTGGGGCTCTGCCTCGAAGAGCTGAACATCAGATTGGAGAGCTC	1181
DB 164176	CTGGAAGACGTGTTTGTGGGGCTCTGCCTCGAAGAGCTGAACATCAGATTGGAGAGCTC	164235
QY 1182	CACCTCCACGCGACCTTTTTCAGGGGGCTTACGCTTCTCCGTATGCTCTTTCAGGAG	1241
DB 164236	CACCTCCACGCGACCTTTTTCAGGGGGCTTACGCTTCTCCGTATGCTCTTTCAGGAG	164295
QY 1242	ATCGTGGCCTGCCACTTTCATCAAGCCTCGGACTCTCTTGGACTACTGGCAGGCTCTAGAG	1301
DB 164296	ATCGTGGCCTGCCACTTTCATCAAGCCTCGGACTCTCTTGGACTACTGGCAGGCTCTAGAG	164355
QY 1302	AATTCGGGGGGAAGATTGTCGCCCTGTCTGAGGGAGCCAGAGGCACATCCGGACAA	1361
DB 164356	AATTCGGGGGGAAGATTGTCGCCCTGTCTGAGGGAGCCAGAGGCACATCCGGACAA	164415
QY 1362	GTTTCAGATAAACCGTGGGATAGTTTTCGTAGATTTTGGAAGAGGGGGGGGACAGAG	1421
DB 164416	GTTTCAGATAAACCGTGGGATAGTTTTCGTAGATTTTGGAAGAGGGGGGGGACAGAG	164475
QY 1422	GATGCTGTTCTTCAGTGTGAAATCCACGCGAGAAATGCGGTGTTTCATGAAGTCACTGAT	1481
DB 164476	GATGCTGTTCTTCAGTGTGAAATCCACGCGAGAAATGCGGTGTTTCATGAAGTCACTGAT	164535
QY 1482	TAGTTCCTCCACTTGGTCCCGCAGCAATAATAGGCCCTCTCTTGGGCACGACACTCTTC	1541
DB 164536	TAGTTCCTCCACTTGGTCCCGCAGCAATAATAGGCCCTCTCTTGGGCACGACACTCTTC	164595
QY 1542	ATATCAAGTGTGTTGACATACACCTGGATTTTTCGATTTTCAGGGGTTCAGTATCCATGACA	1601
DB 164596	ATATCAAGTGTGTTGACATACACCTGGATTTTTCGATTTTCAGGGGTTCAGTATCCATGACA	164655
QY 1602	TGATGGGTGTACATCTCTAATTTTACAGCAAGGACACAGCAGCTGGGAGGTACAGA	1661
DB 164656	TGATGGGTGTACATCTCTAATTTTACAGCAAGGACACAGCAGCTGGGAGGTACAGA	164715
QY 1662	AACTTGTCCCAAGGCTCACAGCCAGTAGGAGCGGGAATGAAATCGAGCACTGT	1721
DB 164716	AACTTGTCCCAAGGCTCACAGCCAGTAGGAGCGGGAATGAAATCGAGCACTGT	164775
QY 1722	CAGAATCTGGTGGGAGCCCTGACTTGAACCACTCCACGTGCTGCCCTTCAGGAG	1781
DB 164776	CAGAATCTGGTGGGAGCCCTGACTTGAACCACTCCACGTGCTGCCCTTCAGGAG	164835
QY 1782	GGCACTGTATGATGAGTCTCGGAGCGGATCCCTCCATCCCTGTCAGTCCCTCCAC	1841
DB 164836	GGCACTGTATGATGAGTCTCGGAGCGGATCCCTCCATCCCTGTCAGTCCCTCCAC	164895
QY 1842	CTCAGTCCCAAGTCTTGTGCTTTTGGAGCTAAGCCTGGATGACCAAAATTCACCCAC	1901
DB 164896	CTCAGTCCCAAGTCTTGTGCTTTTGGAGCTAAGCCTGGATGACCAAAATTCACCCAC	164955
QY 1902	CTCCTTCATTCAGGGCTGGATGTAGCTGGGATGAGTCAATGTTATCGGCTCGGTACT	1961
DB 164956	CTCCTTCATTCAGGGCTGGATGTAGCTGGGATGAGTCAATGTTATCGGCTCGGTACT	165015
QY 1962	CAACACAACCAAGTTTCATCCAGGAAATGTCCCGCAGTGGATGACATGCTG	2021
DB 165016	CAACACAACCAAGTTTCATCCAGGAAATGTCCCGCAGTGGATGACATGCTG	165075
QY 2022	AGGAACACCCAGCTCGACAGAGTTCTTATAAATGTTAATAGGTTCAGAAACCACT	2081
DB 165076	AGGAACACCCAGCTCGACAGAGTTCTTATAAATGTTAATAGGTTCAGAAACCACT	165135
QY 2082	GCATTCCTGACCTGTACAGACTGCCACACTGCTGACCTGCCCTAGGAGCAGACATC	2141
DB 165136	GCATTCCTGACCTGTACAGACTGCCACACTGCTGACCTGCCCTAGGAGCAGACATC	165195
QY 2142	CCTTCTGAGCCATCTGCTCTCTCATTTTCATCACCCCACTGCTCCCTTTTGTGATC	2201
DB 165196	CCTTCTGAGCCATCTGCTCTCTCATTTTCATCACCCCACTGCTCCCTTTTGTGATC	165255
QY 2202	AATGGGACCCAGCTGCCCGCAGGACACTTTAGGGCTCTCAGTTCAAACTGAAGACA	2261

||||| Db 165256 AATGGGGACAGACACCTGCCCCAGAGAGACATTAGGGTCTCAGTTCAAACTGAAGGACA 165315  
QY 2262 GTTGAACACAGATGGGTTTCATGTGGGATCTGGAGCTTTCTGGAAATCASTTGGAGT 2321  
Db 165316 GTTGAACACAGATGGGTTTCATGTGGGATCTGGAGCTTTCTGGAAATCACTTGGAGT 165375  
QY 2322 CAAGTCAGGATGCTCTCAAGGACCCCTCGGGCTCAGAGCCCTAAAGTGGGCCCTGTGTAA 2381  
Db 165376 CAAGTCAGGATGCTCTCAAGGACCCCTCGGGCTCAGAGCCCTAAAGTGGGCCCTGTGTAA 165435  
QY 2382 GCAGGGTGGTCCGTGCTCCACTTCCCAGAGCTGAGCCCAAGCTCATCTTCATTGAATGTCT 2441  
Db 165436 GCAGGGTGGTCCGTGCTCCACTTCCCAGAGCTGAGCCCAAGCTCATCTTCATTGAATGTCT 165495  
QY 2442 CATTTGGCCGAGGAACAACTGAATCTTCTGGTTTCTGCTTTAGCCCTTCAGTTTGTCTCCGC 2501  
Db 165496 CATTTGGCCGAGGACAACTGAATCTTCTGGTTTCTGCTTTAGCCCTTCAGTTTGTCTCCGC 165555  
QY 2502 TGCTCTCTACCCAGAGGTTTGTGGAGGCTGTGTTGCAGGGTTGTATAAACCAGAGTAC 2561  
Db 165556 TGCTCTCTACCCAGAGGTTTGTGGAGGCTGTGTTGCAGGGTTGTATAAACCAGAGTAC 165615  
QY 2562 TTGCTTAGTTTGGCCCATCAGCCATGGTCAGTGACATGCAAGTAATCTTGTCTCTAA 2621  
Db 165616 TTGCTTAGTTTGGCCCATCAGCCATGGTCAGTGACATGCAAGTAATCTTGTCTCTAA 165675  
QY 2622 TTATAGAATGATTTTCTTTTAAATTTTACTTTTACCAGATTTCATTGCTACACAG 2681  
Db 165676 TTATAGAATGATTTTCTTTTAAATTTTACTTTTACCAGATTTCATTGCTACACAG 165735  
QY 2682 AAGAGCCCTCACATGGCTGTGTCACATATAAATGTTGGACTAAACTCTTA 2731  
Db 165736 AAGAGCCCTCACATGGCTGTGTCACATATAAATGTTGGACTAAACTCTTA 165785  
  
RESULT 5  
E38420  
LOCUS E38420 Novel polypeptide. 10562 bp DNA linear PAT 31-JAN-2002  
DEFINITION E38420  
ACCESSION E38420  
VERSION E38420.1 GI:18626994  
KEYWORDS JP 2000245464-A/2.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Narimatsu, H., Isshiki, S., Togayauchi, A. and Sasaki, K.  
Novel polypeptide  
Patent: JP 2000245464-A 2 12-SEP-2000;  
KIOWA HAKKO KOGYO CO LTD  
OS Homo sapiens (human)  
PN JP 2000245464-A/2  
PD 12-SEP-2000  
PF 25-FEB-1999 JP 1999047571  
PR  
PI HISASHI NARIMATSU, SOICHIRO ISSHIKI, AKIRA TOGAYAUCHI, PI  
KATSUTOSHI SASAKI  
PC C12N15/09, A01K67/027, C12N1/21, C12N5/10, C12N9/10, C12P19/00, PC  
C12P21/02,  
PC C12P21/08, C12Q1/68, G01N33/53// (C12N1/21, C12R1:185), (C12N5/10,  
C12R1:191),  
PC (C12P21/02, C12R1:185), (C12P21/02, C12R1:91), C12N15/00, C12N5/00,  
PC (C12N5/00, C12R1:91)  
CC  
CH Key Location/Qualifiers  
FT Promoter (1)..(5000)  
FT exon (5001)..(5140)  
FT exon (5001)..(5273)  
FT exon (5459)..(5567)  
FT exon (7427)..(7586)  
FT exon (8234)..(10562).

FEATURES source Location/Qualifiers  
1..10562  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 2610 a 2415 c 2574 g 2963 t  
ORIGIN  
Query Match 83.9%; Score 2329; DB 6; Length 10562;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 402 ATGGCTTTCCCAAGATGAGATTGATATATTTCCTCTGTTCTGGGGGCTCTTTGT 461  
Db 8234 ATGGCTTTCCCAAGATGAGATTGATATATTTCCTCTGTTCTGGGGGCTCTTTGT 8293  
QY 462 TTGTATTTTAGCATGTACAGTCTAAATCCTTTTCAAGAACAGTCTTTGTTTACAAGAA 521  
Db 8294 TTGTATTTTAGCATGTACAGTCTAAATCCTTTTCAAGAACAGTCTTTGTTTACAAGAA 8353  
QY 522 GACGGGAATCTTCTTAAGTCTCCAGATACAGAGTGCAGGAGACACCTCCCTCTCTCGTC 581  
Db 8354 GACGGGAATCTTCTTAAGTCTCCAGATACAGATGCAGGAGACACCTCCCTCTCTCGTC 8413  
QY 582 CTGCTGTGACCTCATCCCAACAAACAGTTGGCTGAGCGCATGGCCATCCGCGACAGCTGG 641  
Db 8414 CTGCTGTGACCTCATCCCAACAAACAGTTGGCTGAGCGCATGGCCATCCGCGACAGCTGG 8473  
QY 642 GGGAAAGAGAGATGGTGAAGGGAAGAGAGCTGAAGACATCTTCTCTGGGACACC 701  
Db 8474 GGGAAAGAGAGATGGTGAAGGGAAGAGAGCTGAAGACATCTTCTCTGGGACACC 8533  
QY 702 AGCAGTGCAGCGGAACGAAGAGTGGACCAGAGAGCCAGGACAGCGGAGCATTTATC 761  
Db 8534 AGCAGTGCAGCGGAACGAAGAGTGGACCAGAGAGCCAGGACAGCGGAGCATTTATC 8593  
QY 762 CAGAAGGATTTCTAGACGTCTATTACAATCTGACCTGAAGACCATGATGGGATAGAA 821  
Db 8594 CAGAAGGATTTCTAGACGTCTATTACAATCTGACCTGAAGACCATGATGGGATAGAA 8653  
QY 822 TGGTCCATCGCTTTTGTCTCAGCGGGCTTTGTGATGAAACAGACTCAGACATGTTTC 881  
Db 8654 TGGTCCATCGCTTTTGTCTCAGCGGGCTTTGTGATGAAACAGACTCAGACATGTTTC 8713  
QY 882 ATCAATGTTGACTATCTGACTGAATGCTTCTGAAGAAACAGAAACAGAGTTTTC 941  
Db 8714 ATCAATGTTGACTATCTGACTGAATGCTTCTGAAGAAACAGAAACAGAGTTTTC 8773  
QY 942 ACTGGCTTTTGAACACTCAATGAGTTTCCCATCAGGAGCCATTCAGCAAGTGGTTGTC 1001  
Db 8774 ACTGGCTTTTGAACACTCAATGAGTTTCCCATCAGGAGCCATTCAGCAAGTGGTTGTC 8833  
QY 1002 AGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGGCACCAGGCTACGTG 1061  
Db 8834 AGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGGCACCAGGCTACGTG 8893  
QY 1062 TTTTCTGGCGAGCTGGCGAGTCAATGTACAATGTCTCCAAGAGCGTCCCATACATATAA 1121  
Db 8894 TTTTCTGGCGAGCTGGCGAGTCAATGTACAATGTCTCCAAGAGCGTCCCATACATATAA 8953  
QY 1122 CTGGAAGACGTGTTTGTGGGGCTCTGCCCTCGAAAGGCTGAACATCAGATTGGAGAGCTC 1181  
Db 8954 CTGGAAGACGTGTTTGTGGGGCTCTGCCCTCGAAAGGCTGAACATCAGATTGGAGAGCTC 9013  
QY 1182 CACTCCAGCGGACCTTTTTCAGGGGCTTACGGTCTTCCCGTATGCCCTTTCAGGAGG 1241  
Db 9014 CACTCCAGCGGACCTTTTTCAGGGGCTTACGGTCTTCCCGTATGCCCTTTCAGGAGG 9073  
QY 1242 ATCGTGGCCTGCCACTTCATCAAGGCTCGGACTCTCTTGGACTACTGGCAGGCTCTAGAG 1301  
Db 9074 ATCGTGGCCTGCCACTTCATCAAGGCTCGGACTCTCTTGGACTACTGGCAGGCTCTAGAG 9133  
QY 1302 AATTCCCGGGGGAGAGTTGTCGGCTGTCTGAGGGAGCCAGAGGCACATCCGGACAA 1361  
|||||

Db	9134	AATTC	CCCGGGGGAAGATTGTC	CGGCTGTCTGAGGGAGCCGACAGGCACATCCGACAA	9199
Qy	1362	GTTTC	CAGATAA	CCCGTGGGATAGTTTTGCTAGATTTTGGAAAGGGCGGACAGAG	1421
Db	9194	GTTTC	CAGATAA	CCCGTGGGATAGTTTTGCTAGATTTTGGAAAGGGCGGACAGAG	9253
Qy	1422	GATGCTG	TTCTT	CAGTGTGAAATCCACGCCAGAATGTCGGTGTTCATGAAGTCACTGAT	1481
Db	9254	GATGCTG	TTCTT	CAGTGTGAAATCCACGCCAGAATGTCGGTGTTCATGAAGTCACTGAT	9313
Qy	1482	TAGTTT	CCCACTT	GGTGGCCAGCAATAATAGCCCGTCTCTTGGCAGCGCACACTTCTC	1541
Db	9314	TAGTTT	CCCACTT	GGTGGCCAGCAATAATAGCCCGTCTCTTGGCAGCGCACACTTCTC	9373
Qy	1542	ATACT	AAAGTGT	TGACATACACACTGGATTTTGCAATTTTCAGGGGTCAAGTATCCTATGACA	1601
Db	9374	ATACT	AAAGTGT	TGACATACACACTGGATTTTGCAATTTTCAGGGGTCAAGTATCCTATGACA	9433
Qy	1602	TGATGGG	TGT	TACCATCTTAATTTTACAGGCAAGGACACAGACAGCTGCGAGAGGTACAGA	1661
Db	9434	TGATGGG	TGT	TACCATCTTAATTTTACAGGCAAGGACACAGACAGCTGCGAGAGGTACAGA	9493
Qy	1662	AACTTGT	CCCCAAGGCT	TCAGCCAGTAGGCATAGGAGCGGGAATGAAATTCGAGCACTGT	1721
Db	9494	AACTTGT	CCCCAAGGCT	TCAGCCAGTAGGCATAGGAGCGGGAATGAAATTCGAGCACTGT	9553
Qy	1722	CAGAA	CTGGTGGG	CAGCCCTGACTTGAACCACTCCCACGTGCTGCCCTTAGGAGG	1781
Db	9554	CAGAA	CTGGTGGG	CAGCCCTGACTTGAACCACTCCCACGTGCTGCCCTTAGGAGG	9613
Qy	1782	GGACACT	GATGATG	AGGTCTCGGAGCCGGCATCTTCCATCCCTGTCGAGTCCCTCCAC	1841
Db	9614	GGACACT	GATGATG	AGGTCTCGGAGCCGGCATCTTCCATCCCTGTCGAGTCCCTCCAC	9673
Qy	1842	CTCAGCT	CCCCAGT	CCCTCTGTTTTGAGCTAAGCCCTGGGATGACCAATTCACCCCAAG	1901
Db	9674	CTCAGCT	CCCCAGT	CCCTCTGTTTTGAGCTAAGCCCTGGGATGACCAATTCACCCCAAG	9733
Qy	1902	CTCCTT	CATT	CACAGGGTGGATGTAGCTGGGATTCAGTCCATGTTATCGGCTCGGTACT	1961
Db	9734	CTCCTT	CATT	CACAGGGTGGATGTAGCTGGGATTCAGTCCATGTTATCGGCTCGGTACT	9793
Qy	1962	CAACACA	CCCCAAGTTT	CATCGAGGAATGTCCCGGAGTGGATGCACTCACTGCTG	2021
Db	9794	CAACACA	CCCCAAGTTT	CATCGAGGAATGTCCCGGAGTGGATGCACTCACTGCTG	9853
Qy	2022	AGGAAC	CCCCAGTCT	CTGCACAGAGTCTTATAAATGTATAAATAGGCTCAGAAACCACT	2081
Db	9854	AGGAAC	CCCCAGTCT	CTGCACAGAGTCTTATAAATGTATAAATAGGCTCAGAAACCACT	9913
Qy	2082	GCATCT	GACCTGCT	GTACAGACTGCCACACTGCTGACCTGCCTAGGAGCAGACATC	2141
Db	9914	GCATCT	GACCTGCT	GTACAGACTGCCACACTGCTGACCTGCCTAGGAGCAGACATC	9973
Qy	2142	CCTTCT	GAGCCATCT	GCTGCTCTCATTTTCAACCCCAACTGTCCTTTGTTTGTATC	2201
Db	9974	CCTTCT	GAGCCATCT	GCTGCTCTCATTTTCAACCCCAACTGTCCTTTGTTTGTATC	10033
Qy	2202	AATGGGA	CACAGCCAC	TCCCCCAGGAGCACATTTAGGGCTCTCAGTTCAAACTGAAGGACA	2261
Db	10034	AATGGGA	CACAGCCAC	TCCCCCAGGAGCACATTTAGGGCTCTCAGTTCAAACTGAAGGACA	10093
Qy	2262	GTTGA	ACTCAGATGGG	TTTCAATGTTGGGATCTTCTGGGAGCTTTCTGGGAATTCAGTTGGAGT	2321
Db	10094	GTTGA	ACTCAGATGGG	TTTCAATGTTGGGATCTTCTGGGAGCTTTCTGGGAATTCAGTTGGAGT	10153
Qy	2322	CAAGTC	AGGATGCTCT	CAAGGACCCCTCGGCTCAGAGCCCTAAAGTGGGCCCTGGTAA	2381
Db	10154	CAAGTC	AGGATGCTCT	CAAGGACCCCTCGGCTCAGAGCCCTAAAGTGGGCCCTGGTAA	10213
Qy	2382	GCAGGG	TGGTCT	CGGTCCACTTCCAGAGCTGAGCCAGGCTCATCTCATTTGAATGTCT	2441
Db	10214	GCAGGG	TGGTCT	CGGTCCACTTCCAGAGCTGAGCCAGGCTCATCTCATTTGAATGTCT	10273

QY	2442	CATTGGCCGAGGAACAACACTTTGCTGTTTGTGTTTGTCTTGTAGCCTTCAGTTTGTCTCCGC	2501
Db	10274	CATTGGCCGAGGAACAACACTTTGCTGTTTGTGTTTGTCTTGTAGCCTTCAGTTTGTCTCCGC	10333
QY	2502	TGCTCTCTACCCAGAGGTTTGTGGAGCCTGTGTTGCCAGGTTGTATAAAACCAAGGTAC	2561
Db	10334	TGCTCTCTACCCAGAGGTTTGTGGAGCCTGTGTTGCCAGGTTGTATAAAACCAAGGTAC	10393
QY	2562	TTGTTTAGTATTTGCCCCATTGAGCCATGCTCAGCTGACATGCAAGTAATCTTGTCTCTAA	2621
Db	10394	TTGTTTAGTATTTGCCCCATTGAGCCATGCTCAGCTGACATGCAAGTAATCTTGTCTCTAA	10453
QY	2622	TTATAGAAATGATTTTCTTTTAAATTTTACTTTACAGACTTTACTTTGTACTACAG	2681
Db	10454	TTATAGAAATGATTTTCTTTTAAATTTTACTTTACAGACTTTACTTTGTACTACAG	10513
QY	2682	AAGAGGCTCACATGGCTGTGCATATAAATGTTGGACTAAACTCCTT	2730
Db	10514	AAGAGGCTCACATGGCTGTGCATATAAATGTTGGACTAAACTCCTT	10562
RESULT	6		
LOCUS	AF145784	933 bp DNA linear	PRI 14-NOV-1999
DEFINITION	Homo sapiens beta1,3 galactosyltransferase-V (B3GALT5) gene, complete cds.		
ACCESSION	AF145784		
VERSION	AF145784.1	GI:6409192	
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 933)		
TITLE	Zhou, D., Berger, E.G. and Hennet, T.		
	Molecular cloning of a human UDP-galactose-4-epimerase gene encoding an O-linked beta1,3 galactosyltransferase gene		
JOURNAL	core3-elongation enzyme		
MEDLINE	Eur. J. Biochem. 263 (2), 571-576 (1999)		
PUBMED	99337698		
REFERENCE	10406968		
AUTHORS	2 (bases 1 to 933)		
TITLE	Zhou, D. and Hennet, T.		
JOURNAL	Direct Submission		
	Submitted (26-APR-1999); Physiology, University of Zurich, Winterthurerstrasse 190, Zurich 8057, Switzerland		
FEATURES	Location/Qualifiers		
source	1..933		

```

source
1. .933
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22"
1. .933
/gene="B3GALT5"
1. .933
/gene="B3GALT5"
/note="glycosyltransferase"
/codon_start=1
/product="beta1,3 galactosyltransferase-v"
/protein_id="AAF07880.1"
/db_xref="GI:6409193"
/translation="NAPFMRLMYICLLVLGALCYRSMYSILNPFKEOSFVYKKDGNF
LKLPDRCROTPTFLVLLVTSHHKQLAERNAIROTWGRMYVKGKOLATFFLLGTTSS
AAETKEVDQESORHGDI IQKDFLDVYNLTKTMGIEWHFRCQAAFYVKTDSDMF
INVDYITELLKKNNRTTRFTGFLKNEFFIRQPFKSWFVSKSYPDWORYPPFCSGTG
YVSGDVA SQVYNVSKSVYPIKLEDFVFGLCLEKLNIRLEELHSQPTFFPGGLRFSVC
LFRRIVAGCHFIAKPTILLDYNQALENREGDCPPV"
227 a 232 c 241 g 233 t
BASE COUNT

```

Query Match	33.6%	Score 933;	DB 9;	Length 933;
Best Local Similarity	100.0%	Pred. No. 0;		



Matches 933; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	402	ATGCTTTCCGAGATGAGATTGATATATTTCCTTCTGGTCTGGGGCTCTTTGT	461
Db	1	ATGCTTTCCGAGATGAGATTGATATATTTCCTTCTGGTCTGGGGCTCTTTGT	60
QY	462	TTGTATTTAGCATGTACAGTCTAAATCTTCAAGAACAGTCTCTTTGTTTACAAGAAA	521
Db	61	TTGTATTTAGCATGTACAGTCTAAATCTTCAAGAACAGTCTCTTTGTTTACAAGAAA	120
QY	522	GACGGAACTTCCCTTAAGCTCCAGATACAGACTCAGCAGACACCTCCCTTCCTCGTC	581
Db	121	GACGGAACTTCCCTTAAGCTCCAGATACAGACTCAGCAGACACCTCCCTTCCTCGTC	180
QY	582	CTGCTGTGACCTCATCCACAAACAGTTGGCTGAGCGCATGCCATCCGGCAGACGTGG	641
Db	181	CTGCTGTGACCTCATCCACAAACAGTTGGCTGAGCGCATGCCATCCGGCAGACGTGG	240
QY	642	GGGAAGAGAGGATGGTGAAGGGAAGAGCTGAGACATTTCTCTCTGGGGACACC	701
Db	241	GGGAAGAGAGGATGGTGAAGGGAAGAGCTGAGACATTTCTCTCTGGGGACACC	300
QY	702	AGCAGTCACGGGAAGAGAGTGGACAGGAGAGCCAGCAGACACGGGACATTAATC	761
Db	301	AGCAGTCACGGGAAGAGAGTGGACAGGAGAGCCAGCAGACACGGGACATTAATC	360
QY	762	CAGAAGGATTTCTTAGACGTCTATTACAACTGACCTGAACCATGATGGGCATAGAA	821
Db	361	CAGAAGGATTTCTTAGACGTCTATTACAACTGACCTGAACCATGATGGGCATAGAA	420
QY	822	TGGGTCCATCGCTTTTGTCTCAGCGCGCTTTGTGATGAAGAACAGACTCAGACATGTTT	881
Db	421	TGGGTCCATCGCTTTTGTCTCAGCGCGCTTTGTGATGAAGAACAGACTCAGACATGTTT	480
QY	882	ATCAATGTTGACTATCTGACTGAACTGCTTCTGAGAAACAGACACACAGGTTTTTC	941
Db	481	ATCAATGTTGACTATCTGACTGAACTGCTTCTGAGAAACAGACACACAGGTTTTTC	540
QY	942	ACTGGCTCTTTGAACATCAATGAGTTTCCCATCAGGCGCCATTTCAGCAAGTGGTTGTC	1001
Db	541	ACTGGCTCTTTGAACATCAATGAGTTTCCCATCAGGCGCCATTTCAGCAAGTGGTTGTC	600
QY	1002	AGTAAATCTGAATATCCGTGGGACAGGTACCCACCATCTGCTCGGCACCGGCTACGTG	1061
Db	601	AGTAAATCTGAATATCCGTGGGACAGGTACCCACCATCTGCTCGGCACCGGCTACGTG	660
QY	1062	TTTCTGCGACGTGGGAGTCAAGTGTACAAATGTCTCCAGAGCGTCCCATACATATAA	1121
Db	661	TTTCTGCGACGTGGGAGTCAAGTGTACAAATGTCTCCAGAGCGTCCCATACATATAA	720
QY	1122	CTGGAAGACGTGTTTGTGGGCTCTGCTCGAAAGGCTGAACATCAGATTGGAGGAGCTC	1181
Db	721	CTGGAAGACGTGTTTGTGGGCTCTGCTCGAAAGGCTGAACATCAGATTGGAGGAGCTC	780
QY	1182	CACCTCCAGCGACCTTTTTCAGGGGCTTACGCTTCTCGATGCTCTTCAGGAGG	1241
Db	781	CACCTCCAGCGACCTTTTTCAGGGGCTTACGCTTCTCGATGCTCTTCAGGAGG	840
QY	1242	ATCGTGGCTTGCACCTTCATCAAGCCTCGGACTCTCTTTGACTATGCGAGCTCTAGAG	1301
Db	841	ATCGTGGCTTGCACCTTCATCAAGCCTCGGACTCTCTTTGACTATGCGAGCTCTAGAG	900
QY	1302	AATTCCTGGGGGAAGATTGTCGCTCTCTGA	1334
Db	901	AATTCCTGGGGGAAGATTGTCGCTCTCTGA	933

RESULT 7  
AB041416  
LOCUS  
DEFINITION  
ACCESSION

AB041416  
Homo sapiens beta1,3-GalT 5 gene for UDP-Gal:GlcNAc  
beta1,3-galactosyltransferase 5, partial cds.

1576 bp  
linear  
PRI 17-OCT-2000

AB041416.1 GI:7593026			
VERSION	AB041416.1 GI:7593026		
KEYWORDS	Homo sapiens (isolate:#056) DNA.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 1576)		
AUTHORS	Liu, Y. and Saitou, N.		
TITLE	Silver Project		
JOURNAL	Published Only in DataBase (2000)		
REFERENCE	2 (bases 1 to 1576)		
AUTHORS	Liu, Y. and Saitou, N.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-APR-2000) Naruya Saitou, National Institute of Genetics, Laboratory of Evolutionary Genetics, 1111 Yata, Mishima, Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp, URL:http://sayer.lab.nig.ac.jp/~silver/, tel:81-559-81-6790, Fax:81-559-81-6789)		
FEATURES	Location/Qualifiers		
source	1..1576		
	/organism="Homo sapiens"		
	/isolate="#056"		
	/db_xref="taxon:9606"		
	/note="human sequence used for primer design based on Acc# AB020337"		
exon	<1..27		
	/number=3		
intron	28..674		
	/number=3		
gene	675..1576		
	/gene="beta1,3-GalT 5"		
CDS	675..>1576		
	/gene="beta1,3-GalT 5"		
	/codon_start=1		
	/product="UDP-Gal:GlcNAc beta1,3-galactosyltransferase 5"		
	/protein_id="BAA94501.1"		
	/db_xref="GI:7593027"		
	/translation="MAFPMRLMYICLLVLGALCLYFSMSLNPFFKQSFVYKKDGNF LKLPDRCQRPPELVLLVTSRSHKOLAERMAIQTWGKERTVKGKQLKFFLLGTTSS AAETKEVDQESQRHGDIIQKDFLDVYNLTLMGMIEWYHRCFQPAARVPMKTTDSMF INVYLTLLKLNRTTFTGFLKLNFFIROPFFKFWKSKSYEPMDRYPPFCSTGTG YVFGSDVASQVYNVSKSPYIKLEDVFGVGLERLNLRLLEELHSQPTFFPGGLRFSVC LFRIVACHFIKPRLLDLYQWALE"		
exon	675..>1576		
	/gene="beta1,3-GalT 5"		
	/number=4		
BASE COUNT	415 a	373 c	387 g 401 t
ORIGIN			
Query Match	28.8%; Score 800; DB 9; Length 1576;		
Best Local Similarity	99.8%; Pred. No. 0;		
Matches 900; Conservative	0;	Mismatches 2;	Indels 0; Gaps 0;
QY	402	ATGGCTTTCCGAGATGAGATTGATATATTTCCTTCTGGTCTGGGGCTCTTTGT	461
Db	675	ATGGCTTTCCGAGATGAGATTGATATATTTCCTTCTGGTCTGGGGCTCTTTGT	734
QY	462	TTGTATTTAGCATGTACAGTCTAAATCTTTCAAGAACAGTCCCTTTGTTTACAAGAAA	521
Db	735	TTGTATTTAGCATGTACAGTCTAAATCTTTCAAGAACAGTCCCTTTGTTTACAAGAAA	794
QY	522	GACGGGAACCTTCCCTTAAGCTCCAGATACAGACTGAGCGACAGACCTCCCTTCCTCGTC	581
Db	795	GACGGGAACCTTCCCTTAAGCTCCAGATACAGACTGAGCGACAGACCTCCCTTCCTCGTC	854
QY	582	CTGCTGTGACCTCATCCACAAACAGTTGGCTGAGCGCATGCCATCCGGCAGACGTGG	641
Db	855	CTGCTGTGACCTCATCCACAAACAGTTGGCTGAGCGCATGCCATCCGGCAGACGTGG	914
QY	642	GGGAAGAGAGGATGGTGAAGGGAAGAGCGTGAAGACATTTCTTCTCTGGGACACC	701
Db	915	GGGAAGAGAGGCGTGAAGGGAAGAGCGTGAAGACATTTCTTCTCTGGGACACC	974



QY 702 AGCAGTGCAGCGGAACGAAAGAGGTGGACAGGAGCGAGCGACGACGCGGACGACATTATC 761  
Db 975 AGCAGTGCAGCGGAACGAAAGAGGTGGACAGGAGCGAGCGACGACGCGGACGACATTATC 1034  
QY 762 CAGAAAGATTTCTAGACGCTTATTACAATCTGACCTGAAGACCATGATGGGCATAGAA 821  
Db 1035 CAGAAAGATTTCTAGACGCTTATTACAATCTGACCTGAAGACCATGATGGGCATAGAA 1094  
QY 822 TGGGTCCATCGCTTTTGTCTCAGCGGGGCTTTGTGATGAAGAACAGACTCAGACATGTC 881  
Db 1095 TGGGTCCATCGCTTTTGTCTCAGCGGGGCTTTGTGATGAAGAACAGACTCAGACATGTC 1154  
QY 882 ATCAATGTTGACTATCTGACTCACTGCTCTCGAAGAAACAGACAAACAGGTTTTC 941  
Db 1155 ATCAATGTTGACTATCTGACTCACTGCTCTCGAAGAAACAGACAAACAGGTTTTC 1214  
QY 942 ACTGGCTCTTGAAGACTCAATGAGTTTCCATCAGGACGACATTCAGCAAGTGGTTGTC 1001  
Db 1215 ACTGGCTCTTGAAGACTCAATGAGTTTCCATCAGGACGACATTCAGCAAGTGGTTGTC 1274  
QY 1002 AGTAAATCTGAATATCCGTCGAGAGTACCCACCATTCCTGCTCCGCGACCGGTACGTG 1061  
Db 1275 AGTAAATCTGAATATCCGTCGAGAGTACCCACCATTCCTGCTCCGCGACCGGTACGTG 1334  
QY 1062 TTTCTGCGGACGTCGCGAGTCAAGTGTACAATGTCTCAGAGCGGTCCCATACATTAA 1121  
Db 1335 TTTCTGCGGACGTCGCGAGTCAAGTGTACAATGTCTCAGAGCGGTCCCATACATTAA 1394  
QY 1122 CTGGAAGACGTTTGTGGGGCTCTCCCTCGAAGAGCTGAACATCAGATTGGAGAGCTC 1181  
Db 1395 CTGGAAGACGTTTGTGGGGCTCTCCCTCGAAGAGCTGAACATCAGATTGGAGAGCTC 1454  
QY 1182 CACTCCCGACGACCTTTTCCAGGGGCTTACGGTCTCTCCCTATCCCTTTCAGGAGG 1241  
Db 1455 CACTCCCGACGACCTTTTCCAGGGGCTTACGGTCTCTCCCTATCCCTTTCAGGAGG 1514  
QY 1242 ATCGTGGCTGCCACTTCATCAAGCCTCGGACCTCTTGGACACTTGGAGGCTCTAGAG 1301  
Db 1515 ATCGTGGCTGCCACTTCATCAAGCCTCGGACCTCTTGGACACTTGGAGGCTCTAGAG 1574  
QY 1302 AA 1303  
Db 1575 AA 1576  
RESULT 8  
AB041413  
LOCUS AB041413 1565 bp DNA linear PRI 13-APR-2000  
DEFINITION Homo sapiens betal,3-GALT 5 gene for UDP-Gal:GlcNAc  
betal,3-galactosyltransferase 5, partial cds.  
ACCESSION AB041413  
VERSION AB041413.1 GI:7593020  
KEYWORDS  
SOURCE Homo sapiens (isolate:human-NR) DNA.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1565)  
Liu,Y. and Saitou,N.  
Silver Project  
Published Only in Database (2000)  
2 (bases 1 to 1565)  
Liu,Y. and Saitou,N.  
Direct Submission  
Submitted (11-APR-2000) Naruya Saitou, National Institute of  
Genetics, Laboratory of Evolutionary Genetics; 1111 Yata, Mishima,  
Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,  
URL:http://sayer.lab.nig.ac.jp/~silver/, Tel:81-559-81-6790,  
Fax:81-559-81-6789)  
Location/Qualifiers  
1..1565  
/organism="Homo sapiens"  
/isolate="human-NR"

/db\_xref="taxon:9606"  
/notes="human sequence used for primer design based on Acc#  
AB020337"  
<1..27  
28..674  
/number=3  
/number=3  
675..1565  
/gene="betal,3-Galt 5"  
675..1565  
/gene="betal,3-Galt 5"  
/codon\_start=1  
/product="UDP-Gal:GlcNAc betal,3-galactosyltransferase 5"  
/protein\_id="BAA94498.1"  
/db\_xref="GI:7593021"  
/translation="MAPPKMLMYICLLVLGALCLYFSMYSLNPKESQSVYKDGNE  
LKLPTDCRTPPELVLLVTSKQLAERMAIROTWKRXYKQKLLTFFLLGTTSS  
AAETKEVDQESORHGLIOADELDVYLNLLTKWMEIWHRECPQAAVMKTDSDMF  
INVDYELLKRNRTFRFTGFLKNEFPPIRQPFKSFVSKSEYWDYRPFPCSGTG  
YVFGSDVAVQVNVSKVPIKLEDFVGLERLNIHLEELHSPOTTFPPGLRFSVC  
LFRRIVACHFIKPTLTWTR"  
675..1565  
/gene="betal,3-Galt 5"  
/number=4  
BASE COUNT 410 a 370 c 385 g 398 t 2 others  
ORIGIN  
Query Match 27.9%; Score 775; DB 9; Length 1565;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 875; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 402 ATGGCTTCCCGAAGATGAGATTGATATATTTGCCCTTCTGGTTCTGGGGCTCTTTGT 461  
Db 675 ATGGCTTCCCGAAGATGAGATTGATATATTTGCCCTTCTGGTTCTGGGGCTCTTTGT 734  
QY 462 TTGATTTTACGATGTACAGTCTAAATCCCTTCAAGAACAGTCCCTTTTGTACAGAAA 521  
Db 735 TTGATTTTACGATGTACAGTCTAAATCCCTTCAAGAACAGTCCCTTTTGTACAGAAA 794  
QY 522 GACGGAACTTCCCTTAAGCTCCCGATACAGACTGCAGGACAGACCTTCCCTCGTC 581  
Db 795 GACGGAACTTCCCTTAAGCTCCCGATACAGACTGCAGGACAGACCTTCCCTCGTC 854  
QY 582 CTGCTGCTGACCTCATCCCAACAGTTGGCTGAGCGCATGGCCATCGGCGAGAGCTGG 641  
Db 855 CTGCTGCTGACCTCATCCCAACAGTTGGCTGAGCGCATGGCCATCGGCGAGAGCTGG 914  
QY 642 GGGAAAGAGAGGATGGTGAAGGAAAGCAGCTGAAGACATTTCTCTCTCTGGGGACACC 701  
Db 915 GGGAAAGAGAGGATGGTGAAGGAAAGCAGCTGAAGACATTTCTCTCTCTGGGGACACC 974  
QY 702 AGCAGTGCAGCGGAACGAAAGAGGTGGACAGGAGCGACGACGCGGACGACATTATC 761  
Db 975 AGCAGTGCAGCGGAACGAAAGAGGTGGACAGGAGCGACGACGCGGACGACATTATC 1034  
QY 762 CAGAAAGATTTCTAGACGCTTATTACAATCTGACCTGAAGACCATGATGGGCATAGAA 821  
Db 1035 CAGAAAGATTTCTAGACGCTTATTACAATCTGACCTGAAGACCATGATGGGCATAGAA 1094  
QY 822 TGGGTCCATCGCTTTTGTCTCAGCGGGGCTTTGTGATGAAGAACAGACTCAGACATGTC 881  
Db 1095 TGGGTCCATCGCTTTTGTCTCAGCGGGGCTTTGTGATGAAGAACAGACTCAGACATGTC 1154  
QY 882 ATCAATGTTGACTATCTGACTCACTGCTCTCGAAGAAACAGACAAACAGGTTTTC 941  
Db 1155 ATCAATGTTGACTATCTGACTCACTGCTCTCGAAGAAACAGACAAACAGGTTTTC 1214  
QY 942 ACTGGCTCTTGAAGACTCAATGAGTTTCCATCAGGACGACATTCAGCAAGTGGTTGTC 1001  
Db 1215 ACTGGCTCTTGAAGACTCAATGAGTTTCCATCAGGACGACATTCAGCAAGTGGTTGTC 1274  
QY 1002 AGTAAATCTGAATATCCGTCGAGAGTACCCACCATTCCTGCTCCGCGACCGGTACGTG 1061

Db	1275	AGTAAATCTGAATATCCGTGGGACAGGTACCCACCATCTTGCTCCGGCACCGGCTAGCTG	1334
QY	1062	TTTTCTGGGACGTTGGGAGTACAGTGTACAAATGTCTCCAGAGCGCTCCCATACATTA	1121
Db	1335	TTTTCTGGGACGTTGGGAGTACAGTGTACAAATGTCTCCAGAGCGCTCCCATACATTA	1394
QY	1122	CTGGAAGACGTTGTTGGGGCTCTGCCCTCGAAGAGCTGAACATCAGATGGAGGAGCTC	1181
Db	1395	CTGGAAGACGTTGTTGGGGCTCTGCCCTCGAAGAGCTGAACATCAGATGGAGGAGCTC	1454
QY	1182	CACCTCCAGCGGACCTTTTTTCCAGGGGCTTACGCTTCTCCGTATGCCTCTTTCAGGAGG	1241
Db	1455	CACCTCCAGCGGACCTTTTTTCCAGGGGCTTACGCTTCTCCGTATGCCTCTTTCAGGAGG	1514
QY	1242	ATCGTGGCCTGCCACTTCATCAAGGCTCGGACTCTCT	1278
Db	1515	ATCGTGGCCTGCCACTTCATCAAGGCTCGGACTCTCT	1551
RESULT 9			
LOCUS	HS6078	Homo sapiens beta3gal-T6 gene.	933 bp DNA linear PRI 11-MAY-2000
DEFINITION	AJ006078		
ACCESSION	AJ006078.1	GI:7799922	
VERSION			
KEYWORDS		beta-1,3-galactosyltransferase; beta3gal-T6 gene.	
SOURCE		Homo sapiens	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
TITLE		Anado,M., Carneiro,F. and Clausen,H.	
JOURNAL		Cloning and expression of two beta-1,3-galactosyltransferases:	
REFERENCE		beta3gal-T5 and beta3gal-T6	
AUTHORS		Unpublished	
JOURNAL		2 (bases 1 to 933)	
REFERENCE		Anado,M.	
AUTHORS		Direct Submission	
JOURNAL		Submitted (11-MAY-1998) Anado M., Department of Oral Diagnostics,	
FEATURES		Royal dental School, Norre Alle 20, 2200 Copenhagen, DENMARK	
source		Location/Qualifiers	
gene		1..933	/organism="Homo sapiens"
CDS		1..933	/db_xref="taxon:9606"
		1..933	/gene="beta3gal-T6"
		1..933	/gene="beta3gal-T6"
		/codon_start=1	
		/product="beta-1,3-galactosyltransferase"	
		/protein_id="CAB91547.1"	
		/db_xref="GI:7799923"	
		/translation="MAPPKRLMYICLLVLGALCLYYSMSLNPFKEQSFVYKDGNE	
		LKLPDPCDQTPFLVLLVTSKQLAERMAIRQTGWKERTVKGKQLKTFLLGTTSS	
		AAETKEVDQESRQHDIIQKDFLDVYNLLTKMNGIEWHVFPCQAFAFMKTDSDMF	
		INVDYELLKKNRTRFTFLKNEFPIROPEFSKFWSEVPWDRYPPFCSGTG	
		VYFSDVACQVNVYSRVPYIKLEDFVGLCLERLNRLLELHSHOPTFFPGLRFSVC	
		LFRRIVACHIKRFTLLDYWQALENSGEDCPY"	
BASE COUNT	229 a	234 c	240 g
ORIGIN		229 t	1 others
Query Match	25.4%	Score 704;	DB 9; Length 933;
Best Local Similarity	99.6%	Pred. No. 0;	
Matches	834; Conservative	0; Mismatches	3; Indels
		0; Gaps	0;
QY	478	ACAGTCTAAATCTTTCAAGAAGACAGTCCTTTGTTTACAAAGAAACGCGAACTCCCTTA	537
Db	77	ACAGTCTAAATCTTTCAAGAAGACAGTCCTTTGTTTACAAAGAAACGCGAACTCCCTTA	136
QY	538	AGCTCCAGATACAGACTGCAGGACAGACCTCCCTTCTCGTCTGCTGCTGACTCAT	597
Db	137	AGCTCCAGATACAGACTGCAGGACAGACCTCCCTTCTCGTCTGCTGCTGACTCAT	196

QY	598	CCCACAAACAGTTGGCTGAGCGCATGCCCATCCGGCAGACGTGGGGGAAAGAGAGATGG	657
Db	197	CCCACAAACAGTTGGCTGAGCGCATGCCCATCCGGCAGACGTGGGGGAAAGAGAGACGG	256
QY	658	TGAAGGGAAGACAGCTGAAGACATTTCTTCTCTCTGGGGACCAACAGCAGTGCAGCGGAA	717
Db	257	TGAAGGGAAGACAGCTGAAGACATTTCTTCTCTCTGGGGACCAACAGCAGTGCAGCGGAA	316
QY	718	CCAAAGAGGTGGACAGGAGAGCGACGACACGGGGACATTTATCCAGAGGATTTCCCTAG	777
Db	317	CCAAAGAGGTGGACAGGAGAGCGACGACACGGGGACATTTATCCAGAGGATTTCCCTAG	376
QY	778	ACGTCTATTACAAATCTGACCCCTGAAGACCATGATGGGCATAGAATGGGTCCATCCTTT	837
Db	377	ACGTCTATTACAAATCTGACCCCTGAAGACCATGATGGGCATAGAATGGGTCCATCCTTT	436
QY	838	GTCCCTCAGCGCGGTTTGTGATGAAACACAGCTCAGACATGTTTCATCAATGTTGACTATC	897
Db	437	GTCCCTCAGCGCGGTTTGTGATGAAACACAGCTCAGACATGTTTCATCAATGTTGACTATC	496
QY	898	TSAGTGAAGTCTTCTTGAAGAAACAGAACACACCGAGTCTTTCACCTGGCTTCTTGAAC	957
Db	497	TSAGTGAAGTCTTCTTGAAGAAACAGAACACACCGAGTCTTTCACCTGGCTTCTTGAAC	556
QY	958	TCAATGAGTTTCCCATCAGGACGCCATTTCAGCAAGTGGTTTGTGAGTAAATCTGAATATC	1017
Db	557	TCAATGAGTTTCCCATCAGGACGCCATTTCAGCAAGTGGTTTGTGAGTAAATCTGAATATC	616
QY	1018	CGTGGACAGGTACCCACCATCTGCTCCGGCAGCGCTACGTGTTTCTGGGACGTTG	1077
Db	617	CGTGGACAGGTACCCACCATCTGCTCCGGCAGCGCTACGTGTTTCTGGGACGTTG	676
QY	1078	CGAGTCAGGTGACAAATGTCTCCAGAGCGTCCCATACATTAAGTGGAGACGTTGTTG	1137
Db	677	CGAGTCAGGTGACAAATGTCTCCAGAGCGTCCCATACATTAAGTGGAGACGTTGTTG	736
QY	1138	TGGGCTCTGCTCGAAGAGCTGAACATCAGATTGGAGGAGTCCACTCCAGCGGACCT	1197
Db	737	TGGGCTCTGCTCGAAGAGCTGAACATCAGATTGGAGGAGTCCACTCCAGCGGACCT	796
QY	1198	TTTTTCCAGGGGCTTACGCTTCTCCGTATGCCTCTTTCAGGAGGATCGTGGCTGCCACT	1257
Db	797	TTTTTCCAGGGGCTTACGCTTCTCCGTATGCCTCTTTCAGGAGGATCGTGGCTGCCACT	856
QY	1258	TCATCAAGCCTCGGACTCTCTTGGACTACTGCGAGGCTCTAGAGAATTCGCGGGGGAAG	1317
Db	857	TCATCAAGCCTCGGACTCTCTTGGACTACTGCGAGGCTCTAGAGAATTCGCGGGGGAAG	916
QY	1318	ATTGTCGGGCTGCTGA	1334
Db	917	ATTGTCGGGCTGCTGA	933
RESULT 10			
LOCUS	AB041415	1570 bp DNA linear	PRI 13-APR-2000
DEFINITION		Pan paniscus beta1,3-GalT 5 gene for UDP-Gal:GlcNAc	
ACCESSION	AB041415		
VERSION	AB041415.1	GI:7593024	
KEYWORDS		Pan paniscus (isolate:bonobo-05); DNA.	
SOURCE		Pan paniscus	
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
REFERENCE		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.	
AUTHORS		1 (bases 1 to 1570)	
TITLE		Liu,Y. and Saitou,N.	
JOURNAL		Silver Project	
REFERENCE		Published Only in Database (2000)	
AUTHORS		2 (bases 1 to 1570)	
JOURNAL		Liu,Y. and Saitou,N.	
REFERENCE		Direct Submission	
TITLE		Submitted (11-APR-2000) Naruya Saitou, National Institute of	

Genetics, Laboratory of Evolutionary Genetics; 1111 Yata, Mishima, Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp, URL:http://sayer.lab.nig.ac.jp/~silver/, Tel:81-559-81-6790, Fax:81-559-81-6789)

FEATURES		Location/Qualifiers	
Source		1..1570	
exon		/organism="Pan paniscus"	
intron		/isolate="bonobo-05"	
gene		/db_xref="taxon:9597"	
CDS		/note="human sequence used for primer design based on Acc# AB020337"	
		<1..20	
		/number=3	
		21..667	
		/number=3	
		668..1570	
		/gene="betal,3-Galt 5"	
		668..>1570	
		/gene="betal,3-Galt 5"	
		/codon_start=1	
		/product="UDP-Gal:GlcNAc betal,3-galactosyltransferase 5"	
		/protein_id="BAA94500.1"	
		/db_xref="GI:7593025"	
		/translation="MAFPMRLMYVCLLVGLALCLYFSMYSNLNFKESFVYKKGDNF LKLPDTCROTTPFLVLLVTSRSHQLAERMAIRQWTKERTVKGKQLKFTFLGTTSS AAEIKEDQESQRHGDIQKDFLDGYNLTKMMGIEWHVFQCPQAFVNMKTDSMF INVYLTLLKKNRTFRFTGFLKNEFPIQPFKWFVSKSEVPMDRYPPFCSTG YVSGDVASOVYNVSESVPIKLEDFVGLCLERLNIRLELHQSPTFFPGLRFSVC RFRIVACHFIKPRTLDDYWQLEN"	
exon		668..>1570	
		/gene="betal,3-Galt 5"	
		/number=4	
BASE COUNT	408 a	372 c	392 g 398 t
ORIGIN			
Query Match 11.6%; Score 321; DB 9; Length 1570;			
Best Local Similarity 98.8%; Pred. No. 3.3e-176;			
Matches 771; Conservative 0; Mismatches 9; Indels 0; Gaps 0;			
Qy	525	GGGAACCTCTTAAGCTCCAGATACAGCTGCAGGACACCTCCCTTCTCTGCTCTG 584	
Db	791	GGGAACCTCTTAAGCTCCAGATACAGCTGCAGGACACCTCCCTTCTCTGCTCTG 850	
Qy	585	CTGGTGACCTCATCCACAACAGTTGGCTGAGGCGCATGGCCATCCGCGACAGCTGGGG 644	
Db	851	CTGGTGACCTCATCCACAACAGTTGGCTGAGGCGCATGGCCATCCGCGACAGCTGGGG 910	
Qy	645	AAAGAGAGATGGTGAAGGAAGACAGCTGAAGACATCTTCCTCTGGGACACACAGC 704	
Db	911	AAAGAGAGACGGTGAAGGAAGACAGCTGAAGACATCTTCCTCTGGGACACACAGC 970	
Qy	705	AGTGACGCGGAACGAAGAGGTGGACAGGAGCGACGACGCGGAGACATATCCAG 764	
Db	971	AGTGACGCGGAACGAAGAGGTGGACAGGAGCGACGCGGAGACATATCCAG 1030	
Qy	765	AAGGATTTCTAGACGCTATTATCAATCTGACCTGAAGACCATGATGGGCATAGAATGG 824	
Db	1031	AAGGATTTCTAGACGCTATTATCAATCTGACCTGAAGACCATGATGGGCATAGAATGG 1090	
Qy	825	GTCATCGCTTTTGTCTCAGCGGGGTTTGTGATGAACAGACTCAGACATGTTTATC 884	
Db	1091	GTCATCGCTTTTGTCTCAGCGGGGTTTGTGATGAACAGACTCAGACATGTTTATC 1150	
Qy	885	ATGTTGACTATCTGACTGAACGCTTCTGAAGAAAACAGAAACAGAGTTTTCCTACT 944	
Db	1151	ATGTTGACTATCTGACTGAACGCTTCTGAAGAAAACAGAAACAGAGTTTTCCTACT 1210	
Qy	945	GGCTCTTTGAAACTCAATGAGTTTCCCATCAGGACGACATTCACAAAGTGGTTGTGAGT 1004	
Db	1211	GGCTCTTTGAAACTCAATGAGTTTCCCATCAGGACGACATTCAGTAAGTGGTTGTGAGT 1270	
Qy	1005	AAATCTGAATATCCGTGGACAGGTACCCACCATCTCTCCGCGACCGGTACGTGTTT 1064	

Db	1271	AAATCTGAATATCCGTGGACAGGTACCGCGCATTCGTCCGGCACCAGCTACGTGTTT 1330	
Qy	1065	TCTGCCAGCGTGGCGAGTCAGGTGTACAATGTCTCCAAGAGCGTCCCATACATTAAACTG 1124	
Db	1331	TCCGGCGACGTGGCGAGTCAGGTGTACAATGTCTCCGAGAGCGTCCCATACATTAAACTG 1390	
Qy	1125	GAAGACGTGTTTGTGGGCTCTGCCTCGAAAGCTGAACATCAGATTGGAGGAGCTCCAC 1184	
Db	1391	GAAGACGTGTTTGTGGGCTCTGCCTCGAAAGCTGAACATCAGATTGGAGGAGCTCCAC 1450	
Qy	1185	TCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCTCGTATGCTCTTCAGGAGGATC 1244	
Db	1451	TCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCTCGTATGCTCTTCAGGAGGATC 1510	
Qy	1245	GTGGCTTGCACCTTCATCAAGCTCGGACTCTCTTGAGCTACTGGCAGGCTCTAGAGAAT 1304	
Db	1511	GTGGCTTGCACCTTCATCAAGCTCGGACTCTCTTGAGCTACTGGCAGGCTCTAGAGAAT 1570	
RESULT 11			
AB041414			
LOCUS	AB041414	1566 bp	DNA linear PRI 13-APR-2000
DEFINITION	Pan troglodytes betal,3-Galt 5 gene for UDP-Gal:GlcNAc betal,3-galactosyltransferase 5, partial cds.		
ACCESSION	AB041414		
VERSION	AB041414.1	GI:7593022	
KEYWORDS	Pan troglodytes (isolate:chimp-202) DNA.		
SOURCE	Pan troglodytes		
ORGANISM	Pan troglodytes		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.		
AUTHORS	1 (bases 1 to 1566)		
TITLE	Liu, Y. and Saitou, N.		
JOURNAL	Silver Project		
REFERENCE	Published Only in DataBase (2000)		
AUTHORS	2 (bases 1 to 1566)		
TITLE	Liu, Y. and Saitou, N.		
JOURNAL	Direct Submission		
FEATURES	Submitted (11-APR-2000) Naruya Saitou, National Institute of Genetics, Laboratory of Evolutionary Genetics; 1111 Yata, Mishima, Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp, URL:http://sayer.lab.nig.ac.jp/~silver/, Tel:81-559-81-6790, Fax:81-559-81-6789)		
Source	Location/Qualifiers		
exon	1..1566		
intron	/organism="Pan troglodytes"		
gene	/isolate="chimp-202"		
CDS	/db_xref="taxon:9598"		
	/note="human sequence used for primer design based on Acc# AB020337"		
	<1..28		
	/number=3		
	29..675		
	/number=3		
	675..1566		
	/gene="betal,3-Galt 5"		
	675..>1566		
	/gene="betal,3-Galt 5"		
	/codon_start=1		
	/product="UDP-Gal:GlcNAc betal,3-galactosyltransferase 5"		
	/protein_id="BAA94499.1"		
	/db_xref="GI:7593023"		
	/translation="MAFPMRLMYVCLLVGLALCLYFSMYSNLNFKESFVYKKGDNF LKLPDTCROTTPFLVLLVTSRSHQLAERMAIRQWTKERTVKGKQLKFTFLGTTSS AAEIKEDQESQRHGDIQKDFLDGYNLTKMMGIEWHVFQCPQAFVNMKTDSMF INVYLTLLKKNRTFRFTGFLKNEFPIQPFKWFVSKSEVPMDRYPPFCSTG YVSGDVASOVYNVSESVPIKLEDFVGLCLERLNIRLELHQSPTFFPGLRFSVC RFRIVACHFIKPRTLDDYWQ"		
exon	675..>1566		
	/gene="betal,3-Galt 5"		
	/number=4		
BASE COUNT	404 a	374 c	394 g 393 t 1 others
ORIGIN			

	Query Match	9.1%	Score 252;	DB 9;	Length 1566;
	Best Local Similarity	98.7%;	Pred. No. le-135;		
	Matches 702;	Conservative 0;	Mismatches 9;	Indels 0;	Gaps 0;
QY	491	TTTCAAAGACAGTCCCTTTGTTTACAAAGAACACGGGAACCTTCCTTAAAGTCCOCAGATAC	550		
Db	764	TTTCAAAGAACAGTCCCTTTGTTTACAAAGAACACGGGAACCTTCCTTAAAGTCCOCAGATAC	823		
QY	551	AGACTCGAGCAGACACCTCCCTTCCTCGTCTGCTGCTGCTGCTGCTCATCCACAAACAGTT	610		
Db	824	AGACTCGAGCAGACACCTCCCTTCCTCGTCTGCTGCTGCTGCTCATCCACACACAGTT	883		
QY	611	GGCTGAGCGCATGGCCATCCGGCAGACGCTGGGGGAAAGAGAGATCGTGAAGGGAAGCA	670		
Db	884	GGCTGAGCGCATGGCCATCCGGCAGACGCTGGGGGAAAGAGAGACCGTGAAGGGAAGCA	943		
QY	671	GCTGAAGACATTTCTTCCTCGTGGGACACCAGCAGTGCAGCGGAAACGAAAGAGGTGGA	730		
Db	944	GCTGAAGACATTTCTTCCTCGTGGGACACCAGCAGTGCAGCGGAAACGAAAGAGGTGGA	1003		
QY	731	CCAGGAGAGCCAGCAGACCGGGGACATTTACAGAAGGATTTCTTAGACGCTATTACAA	790		
Db	1004	CCAGGAGAGCCAGCAGACCGGGGACATCATCAGAAGGATTTCTTGACGCTATTACAA	1063		
QY	791	TCTGACCTTGAAGACCATGATGGGCATAGATGGGTCCATCGCTTTTGCTCAGGCGGC	850		
Db	1064	TCTGACCTTGAAGACCATGATGGGCATAGATGGGTCCATCGCTTTTGCTCAGGCGGC	1123		
QY	851	GTTTGTGATGAACACAGACTCAGACATGTTTCATCAATGTTGACTATCTCAGTGAAGTCT	910		
Db	1124	GTTTGTGATGAACACAGACTCAGACATGTTTCATCAATGTTGACTATCTCAGTGAAGTCT	1183		
QY	911	TCTGAAGAAAAACAGACAACCCAGGTTTTTTCAGTGGCTTCTTGAAACTCAATCAGTTTC	970		
Db	1184	TCTGAAGAAAAACAGACAACCCAGGTTTTTTCAGTGGCTTCTTGAAACTCAATCAGTTTC	1243		
QY	971	CATCAGGCAGCCATTTCAGCAAGTGGTTGTCAAGTAAATCTGAAATTCOCGTGGGACAGGTA	1030		
Db	1244	CATCAGGCAGCCATTTCAGTAAAGTGGTTGTCAAGTAAATCTGAAATTCOCGTGGGACAGGTA	1303		
QY	1031	CCACCATTTCTGCTCCGGCACCGGCTACGTTTTTCTTGCGCAGCTGGCAGTCAGGTGTA	1090		
Db	1304	CCCGCATTTCTGCTCCGGCACCGGCTACGTTTTTCTTGCGGACGCTGGCAGTCAGGTGTA	1363		
QY	1091	CAATGTCTCCAGAGCGTCCCATACATTAAACTGGAAGACGTTTTGTGGGGCTCTGCCT	1150		
Db	1364	CAATGTCTCCGAGCGTCCCATACATTAAACTGGAAGACGTTTTGTGGGGCTCTGCCT	1423		
QY	1151	CGAAGGCTGAACATCAGATTGGAGGAGCTCCACTCCCGACCGACGTTTTT	1201		
Db	1424	CGAAGGCTGAACATCAGATTGGAGGAGCTCCACTCCCGACCGACGTTTTT	1474		

RESULT	12
LOCUS	AB041412
DEFINITION	AB041412 linear PRI 17-OCT-2000
ACCESSION	Gorilla gorilla beta1.3-GALT 5 gene for UDP-Gal:GlcNAc
VERSION	beta1.3-galactosyltransferase 5, partial cds.
KEYWORDS	AB041412 GI:7593018
SOURCE	Gorilla gorilla (isolate:#085) DNA.
ORGANISM	Gorilla gorilla Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
REFERENCE	1 (bases 1 to 1360)
AUTHORS	Liu,Y. and Saitou,N.
TITLE	Silver Project
JOURNAL	Published Only in DataBase (2000)
REFERENCE	2 (bases 1 to 1360)
AUTHORS	Liu,Y. and Saitou,N.
TITLE	Direct Submission

**JOURNAL**  
Submitted (11-APR-2000) Naruya Saitou, National Institute of  
Genetics, Laboratory of Evolutionary Genetics, 1111 Yata, Mishima,  
Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,  
URL: <http://sayer.lab.nig.ac.jp/~silver/>, Tel: 81-553-81-6790,  
Fax: 81-559-81-6789)  
**FEATURES**  
Location/Qualifiers  
1. 1360  
/organism="Gorilla gorilla"  
/isolate="#085"  
/db\_xref="taxon:9593"  
/note="human sequence used for primer design based on Acc#

```

FEATURES
source
Location/Qualifiers
1. .1360
/organism="Gorilla gorilla"
/isolate="#085"
/db_xref="taxon:9593"
/note="human sequence used for primer design based on Acc#
AB020337"
<1. .466
/number=3
467. .1360
/genes="betal,3-Galt 5"
467. .>1360
/genes="betal,3-Galt 5"
/codon_start=1
/product="UDP-Gal:GlcNAc betal,3-galactosyltransferase 5"
/protein_id="BAA94497.1"
/db_xref="GI:7593019"
/translation="MAFPKRLMYICLLVLGALCLFSYSLNPKEQSFVYKKDYNF
LKLPTDCQTPPEFLVLTSSHQLAERAIROTGWKERTVKGQLKTFEILGTTSS
RAETVEQDSERHGLDIQKDFLDVYNLILKTMMGIEWHRCFQAAPVMTKDSMF
INVDYLLLLAKNRITRFETFLKLNFEPIRQTSKFWKSKEYEWDRIYPPFCSTG
YFSGDVAQVNVNSESVPYIKLEDFVEVGLCERLINRLEELHSQTPFPFGGLRFVSC
467. .>1360
/genes="betal,3-Galt 5"
/number=4
334 a 331 c 341 g 351 t 3 others
exon
intron
gene
CDS

```

Query Match	8.3%;	Score 230;	DB 9;	Length 1360;
Best Local Similarity	99.4%;	Pred. No. 8.5e-123;		
Matches 330;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	657	GTGAAGGGAACGAGCTGAGACATTCTTCTCTGTGGGACACCAGCAGTGCACGGGAA	716	
Db	722	GTGAAGGGAACGAGCTGAGACATTCTTCTCTGTGGGACACCAGCAGTGCACGGGAA	781	
QY	717	ACGAAGAGGTTGACACAGAGAGCGGACACCGGGACATTATCCAGAAGGATTTCCTA	776	
Db	782	ACGAAGAGGTTGACACAGAGAGCGGACACCGGGACATTATCCAGAAGGATTTCCTA	841	
QY	777	GACGTCTATTACAATCTGACCTTGAGACCATGATGGGCATAGAAATGGTCCATCGCTTT	836	
Db	842	GAYGCTCTATTACAATCTGACCTTGAGACCATGATGGGCATAGAAATGGTCCATCGCTTT	901	
QY	837	TGTCCTCAGGCGGCGTTTGTGATGAACAACAGACTCAGACATGTTTCATCAATGTTGACTAT	896	
Db	902	TGTCCTCAGGCGGCGTTTGTGATGAACAACAGACTCAGACATGTTTCATCAATGTTGACTAT	961	
QY	897	CTGACTGAACCTGCTTCTGAAGAAAAACAGAACCAACCAGGTTTTTTCACGTGGCTTCTTTGAAA	956	
Db	962	CTGACTGAACCTGCTTCTGAAGAAAAACAGAACCAACCAGGTTTTTTCACGTGGCTTCTTTGAAA	1021	
QY	957	CTCAATGAGTTTCCCATCAGGCAGCCATTCCAG	988	
Db	1022	CTCAATGAGTTTCCCATCAGGCAGCCATTCCAG	1053	

[illegible]

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pongo.  
1 (bases 1 to 1579)  
Li.Y. and Saitou.N.  
Silver Project  
Published Only in DataBase (2000)  
2 (bases 1 to 1579)  
Li.Y. and Saitou.N.  
Direct Submission  
Submitted (11-APR-2000) Naruya Saitou, National Institute of Genetics, Laboratory of Evolutionary Genetics, 1111 Yata, Mishima, Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp, URL:htp://sayer.lab.nig.ac.jp/~silver/, Tel:81-559-81-6790, Fax:81-559-81-6789)

FEATURES  
Location/Qualifiers  
1..1579  
/organism="Pongo pygmaeus"  
/isolate="Oran-Pol7"  
/db\_xref="taxon:9600"  
/note="human sequence used for primer design based on Acc# AB020337"  
<1..26  
/number=3  
27..673  
/number=3  
674..1579  
/gene="betal.3-Galt 5"  
674..>1579  
/gene="betal.3-Galt 5"  
/codon\_start=1  
/product="UDP-Gal:GlcNAc betal.3-galactosyltransferase 5"  
/protein\_id="BAA94502.1"  
/db\_xref="GI:7593029"  
exon  
intron  
gene  
CDS  
5.38; Score 147; DB 9; Length 1579;  
Best Local Similarity 100.0%; Pred. No. 4.4e-74;  
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 819 GAATGGTCCATCGCTTTGTCCTCAGCGCGGTTGTGATGAAACAGACTCAGACATG 878  
DB 1094 GAATGGTCCATCGCTTTGTCCTCAGCGCGGTTGTGATGAAACAGACTCAGACATG 1153  
QY 879 TTCATCAATGTTGACTATCTGACTGAATGCTTCTGAGAGAAAACAGACACACCGGTTT 938  
DB 1154 TTCATCAATGTTGACTATCTGACTGAATGCTTCTGAGAGAAAACAGACACACCGGTTT 1213  
QY 939 TTCCTGCTCTTCTGAACTCAATGAG 965  
DB 1214 TTCCTGCTCTTCTGAACTCAATGAG 1240

RESULT 14  
HSM800208  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1746 bp mRNA linear PRI 18-FEB-2000  
Homo sapiens mRNA; cDNA DKFZp586B211 (from clone DKFZp586B211).  
AL049423.1 GI:4500204

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
1 (bases 1 to 1746)  
Wambutt.R., Heubner.D., Mewes.H.W., Gassenhuber.J. and Wiemann.S.  
Direct Submission  
Submitted (10-MAR-1999) MIPS, Am Klopferspitz 18a, D-82152 Martinsried, GERMANY  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
consortium of the German genome project.  
This clone (DKFZp586B211) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://www.mips.biochem.mpg.de/proj/cDNA/.

FEATURES  
Location/Qualifiers  
1..1746  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="524.0 cR from top of Chr3 linkage group"  
/clone="DKFZp586B211"  
/tissue\_type="uterus"  
/clone\_lib="586 (synonym: hutel). Vector pSport1; host DH10B; sites NotI + SalI/MluI"  
/dev\_stage="adult"  
1668..1673  
polyA\_signal  
polyA\_site  
BASE COUNT 577 a 311 c 293 g 565 t  
ORIGIN  
Query Match 1.9%; Score 53; DB 9; Length 1746;  
Best Local Similarity 100.0%; Pred. No. 6.5e-19;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2723 AAACCTCTTAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAAT 2775  
DB 1682 AAACCTTAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 1734  
RESULT 15  
BC013407  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
2771 bp mRNA linear PRI 08-NOV-2001  
Homo sapiens, clone MGC:4518 IMAGE:2988445, mRNA, complete cds.  
BC013407.1 GI:16807136

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REMARK  
COMMENT  
NIH-MGC Project URL: http://mgc.nci.nih.gov  
Contact: MGC help desk  
Email: cgapps-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Institute for Systems Biology  
http://www.systemsbio.org  
contact: amadan@systemsbiology.org  
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Series: IRAL Plate: 10 Row: k Column: 10  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 13376787.

```
FEATURES
  source
    Location/Qualifiers
      1..2771
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="MGC:4518 IMAGE:2988445"
        /tissue_type="Colon, adenocarcinoma"
        /clone_lib="NIH_MGC_15"
        /lab_host="DH10B-R"
        /note="Vector: pOTB7"
        1600..2445
          CDS
            /codon_start=1
            /product="Unknown (protein for MGC:4518)"
            /protein_id="AAH13407.1"
            /db_xref="GI:16807137"
            /translation="MLRFVQKRGNSTVYEWRTGTPTSPSVRPHLEELPEQVAEDAIDW
            GDFGEAVASEGTDGISAAGIDWGIFFPSDSKDPGGDIDGDDAVALQITVLEAG
            TOAPEGVARGPDALTLLLEYTETRNQFLDELMELFILAQRAVELSEEDVLSVSQFQL
            APAILQGGTKERKMTWYSVLEDLIGKLTSLQLQLHFLMILASPRYVDRTVEFLQQKLKQ
            SLLALKKELMVQKQEALEEQAALEPKLDLLLEKTLKLEADISKRYSGRPVNL
            MGTSL"
      BASE COUNT      683 a      647 c      820 g      621 t
      ORIGIN
        Query Match      1.9%; Score 53; DB 9; Length 2771;
        Best Local Similarity 100.0%; Pred. No. 6.5e-19;
        Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
        QY 2723 AACTCTTAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2775
           |||||
        Db 2708 AACTCTTAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2760
           |||||
        Search completed: April 12, 2003, 07:16:45
        Job time : 5888.03 secs
```

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 22:38:14 ; Search time 455.571 Seconds  
(without alignments)  
13717.504 Million cell updates/sec

Title: US-09-914-152-2  
Perfect score: 2775  
Sequence: 1 gtgaattctcttctctgc.....aaaaaaaaaaaaaaaaaaaaa 2775

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0 4370478

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_101002:\*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*
- 24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2775	100.0	2775	21 AAA93875	Human beta-1,3 gal
2	2329	83.9	10562	21 AAA93876	Human beta3Gal-T5
3	798	28.8	1011	21 AAA27959	Human beta3Gal-T5
4	60	2.2	60	24 ABN47639	Human spliced tran
5	52	1.9	861	16 AAQ99552	Thrombopoietin cod
6	51	1.8	190	23 ABV54702	Human prostate exp
7	51	1.8	390	23 ABV55403	Human prostate exp
8	51	1.8	422	23 ABV54507	Human prostate exp
9	50	1.8	215	23 ABV19654	Human prostate exp

10	50	1.8	341	23 ABV49421	Human prostate exp
11	50	1.8	383	22 AA188530	Human polynucleoti
12	50	1.8	386	22 AA183204	Human polynucleoti
13	50	1.8	419	22 AA188760	Human polynucleoti
14	50	1.8	427	22 AA184653	Human polynucleoti
15	50	1.8	430	23 ABV57740	Human prostate exp
16	50	1.8	441	22 AA182363	Human polynucleoti
17	50	1.8	935	24 AAD33050	Human polynucleoti
18	50	1.8	1092	22 AAH34069	Human colon cancer
19	50	1.8	1443	21 AAA26291	Human secreted pro
20	50	1.8	1474	18 AA790174	Human secreted pro
21	50	1.8	1586	14 AAQ47967	Oil seed rape cyst
22	50	1.8	1801	22 AAS26060	Rape acyl-ACP thio
23	50	1.8	1924	24 AB190597	Human cDNA encodin
24	50	1.8	1939	21 AAD02321	Human polynucleoti
25	50	1.8	1939	24 AAD02867	Human serine prote
26	50	1.8	2301	20 AAZ00344	Human serine prote
27	50	1.8	2301	22 AAF89002	Nucleotide sequenc
28	50	1.8	2710	20 AAZ00368	Murine FATP4 codin
29	50	1.8	2710	20 AAZ00355	Nucleotide sequenc
30	50	1.8	2710	22 AAF89017	Murine FATP4 codin
31	50	1.8	2710	22 AAF89030	Murine FATP4 codin
32	50	1.8	8079	24 ABU92313	Chemically treated
33	50	1.8	13919	24 ABU92304	Chemically treated
34	49	1.8	86	22 AAS23317	Human prostate can
35	49	1.8	173	23 ABV57958	Human prostate exp
36	49	1.8	289	23 ABV49828	Human prostate exp
37	49	1.8	324	23 ABV57099	Human prostate exp
38	49	1.8	346	22 AA187483	Human polynucleoti
39	49	1.8	357	22 AA187170	Human polynucleoti
40	49	1.8	365	22 AA187543	Human polynucleoti
41	49	1.8	374	22 AA191055	Human polynucleoti
42	49	1.8	384	23 AA461128	Human polynucleoti
43	49	1.8	392	22 AA187456	Human prostate exp
44	49	1.8	405	22 AA187598	Human polynucleoti
45	49	1.8	406	22 AA187349	Human polynucleoti

ALIGNMENTS

RESULT 1  
AAA93875  
ID AAA93875 standard; DNA; 2775 BP.  
AC AAA93875;  
XX  
XX 15-JAN-2001 (first entry)  
DE Human beta-1,3 galactose transferase encoding DNA.  
KW Beta-1,3 galactose transferase; treatment; diagnosis; cancer; human;  
KW digestive system; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO2000050608-A1.  
XX  
XX 31-AUG-2000.  
XX  
XX 24-FEB-2000; 2000WO-JP01070.  
XX  
XX 25-FEB-1999; 99JP-0047571.  
XX  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
XX  
XX Narimatsu H, Isshiki S, Togayachi A, Sasaki K;  
XX  
XX WPI; 2000-549409/50.  
XX  
XX P-PSDB; AAB93875.  
XX  
XX Beta-1,3 galactose transferase and DNA encoding it, useful for  
PT synthesis of type 1 sialyl Lewis, a carbohydrate for treatment of

PT digestive system cancer -  
XX Claim 5; Page 99-102; 123pp; Japanese.  
XX This invention relates to a polypeptide (I) with beta-1,3 galactose  
CC transferase activity, or variants of (I) comprising amino acid additions,  
CC deletions and/or substitutions. Included in the invention is DNA encoding  
CC all or part of (I); expression vectors containing the DNA, host cells  
CC transformed by the vectors; a method for the preparation of the  
CC polypeptide by culture of the transformants or by expression in the milk  
CC of a transgenic mammal, and antibodies recognising (I). The Beta-1,3  
CC galactose transferase protein transfers galactose by beta-1,3 bonding to  
CC N-acetylglucosamine present in a non-cyclic carbohydrate chain (such as  
CC GlcNAc-beta1-3gal-beta1-4Glc) to give Gal-beta1-3GlcNAc. The protein and  
CC DNA encoding it are useful for the treatment and diagnosis of cancer of  
CC the digestive system. The present sequence represents Beta-1,3 galactose  
CC transferase encoding DNA.  
XX Sequence 2775 BP; 581 A; 598 C; 569 G; 727 T; 0 other;  
SQ

Query Match 100.0%; Score 2775; DB 21; Length 2775;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2775; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGAATTCCTCTTCTGCTGGAGCTGGGATATCTTCTCTCTGCTGCGCTTGGACATCA 60  
DB 1 GTGAATTCCTCTTCTGCTGGAGCTGGGATATCTTCTCTCTGCTGCGCTTGGACATCA 60  
QY 61 GAGCTCAGGCTCTGCGCTTTGGACCGAGGATTTATCAACAGAGTTTCTGGGTTTC 120  
DB 61 GAGCTCAGGCTCTGCGCTTTGGACCGAGGATTTATCAACAGAGTTTCTGGGTTTC 120  
QY 121 TCAGGCTTTGGGCTTGGACTGATAGTTACACCATTTGSCATATCTGTTCTGAGGCTTT 180  
DB 121 TCAGGCTTTGGGCTTGGACTGATAGTTACACCATTTGSCATATCTGTTCTGAGGCTTT 180  
QY 181 GGTCTGGAGTACAGCCACATCTCGGATCCAGGCTCCAGCTTGCATGTCGCTGTCTAC 240  
DB 181 GGTCTGGAGTACAGCCACATCTCGGATCCAGGCTCCAGCTTGCATGTCGCTGTCTAC 240  
QY 241 GTGATTCCTGTCAAGTACCATTTTGGTAAACAAACCAAGCCAGACCTGTGATAATTA 300  
DB 241 GTGATTCCTGTCAAGTACCATTTTGGTAAACAAACCAAGCCAGACCTGTGATAATTA 300  
QY 301 TGGAGCATTTACACTGACAGTCTTTTGAGACAAATTTCCCTCTTGGCATTTACACTGTGG 360  
DB 301 TGGAGCATTTACACTGACAGTCTTTTGAGACAAATTTCCCTCTTGGCATTTACACTGTGG 360  
QY 361 CTATTAGCTTCAACACAGAGTTCCTCTTACCCAGCAAAAATGCTTTCCCGAAGATGA 420  
DB 361 CTATTAGCTTCAACACAGAGTTCCTCTTACCCAGCAAAAATGCTTTCCCGAAGATGA 420  
QY 421 GATTGATGATATTTGGCTTCTGCTGCTGGGGCTCTTTGTTGTTATTTAGCATGTACA 480  
DB 421 GATTGATGATATTTGGCTTCTGCTGCTGGGGCTCTTTGTTGTTATTTAGCATGTACA 480  
QY 481 GTCTAAATCCTTTCAAGAACAGTCCCTTTGTTTACAAGAAAGACGGAACTTCCCTTAAGC 540  
DB 481 GTCTAAATCCTTTCAAGAACAGTCCCTTTGTTTACAAGAAAGACGGAACTTCCCTTAAGC 540  
QY 541 TCCAGATACAGACTGACAGGACACATCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
DB 541 TCCAGATACAGACTGACAGGACACATCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
QY 601 ACAACAGTTGGCTGAGCGATGCCATCCCGCAGACCTGGGGAAAGAGAGATGGTGA 660  
DB 601 ACAACAGTTGGCTGAGCGATGCCATCCCGCAGACCTGGGGAAAGAGAGATGGTGA 660  
QY 661 AGGAAAGCAGCTGAAGACATTTCTCCCTGGGGACACACAGCAGTGCAGCGGAAACGA 720  
DB 661 AGGAAAGCAGCTGAAGACATTTCTCCCTGGGGACACACAGCAGTGCAGCGGAAACGA 720  
QY 721 AAGAGTGGACAGGAGGCGACAGCGGAGCATTTACAGAAAGGATTTCTCTAGACG 780

DB 721 AAGAGTGGACAGGAGCGGACAGCGGAGCATTTATCCAGAAAGGATTTCTCTAGACG 780  
QY 781 TCTATTACAATCTGACCTGAAGACCATGATGGGATGAATGGTCCATCGCTTTTGTGTC 840  
DB 781 TCTATTACAATCTGACCTGAAGACCATGATGGGATGAATGGTCCATCGCTTTTGTGTC 840  
QY 841 CTCAGCGGCTTTGTGATGAAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGA 900  
DB 841 CTCAGCGGCTTTGTGATGAAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGA 900  
QY 901 CTGAACCTGCTTCTGAAGAAAAACAGAAACACAGGTTTTTCTACTGGCTTCTTGAACCTCA 960  
DB 901 CTGAACCTGCTTCTGAAGAAAAACAGAAACACAGGTTTTTCTACTGGCTTCTTGAACCTCA 960  
QY 961 ATGAGTTTCCCATCAGGCGCATTTCAAGCAAGTGGTTTGTGCTAGTAATCTGAATATCCGT 1020  
DB 961 ATGAGTTTCCCATCAGGCGCATTTCAAGCAAGTGGTTTGTGCTAGTAATCTGAATATCCGT 1020  
QY 1021 GGGACAGGTACCCACCATCTCTCGGACCGGCTACGTTTCTTGGCGAGCTGGCGA 1080  
DB 1021 GGGACAGGTACCCACCATCTCTCGGACCGGCTACGTTTCTTGGCGAGCTGGCGA 1080  
QY 1081 GTCAGGTGTACAATGCTCTCAAGAGCGTCCCATACATTTAACTTGAAGAGCGTGTGTGG 1140  
DB 1081 GTCAGGTGTACAATGCTCTCAAGAGCGTCCCATACATTTAACTTGAAGAGCGTGTGTGG 1140  
QY 1141 GGCTCTGCTCGAAAGGCTGAACATCAGATTGAGAGCTCCACTCCACGCGACCTTTT 1200  
DB 1141 GGCTCTGCTCGAAAGGCTGAACATCAGATTGAGAGCTCCACTCCACGCGACCTTTT 1200  
QY 1201 TTCAGGCGGCTTACGTTTCTCCGTATGCTCTTTCAGGAGGATCGTGGCTTGGCACTTCA 1260  
DB 1201 TTCAGGCGGCTTACGTTTCTCCGTATGCTCTTTCAGGAGGATCGTGGCTTGGCACTTCA 1260  
QY 1261 TCAAGCTTCGCACTCTCTTGGACTACTTGGAGCTCTAGAGAATTCCTCCGGGGGAAGATT 1320  
DB 1261 TCAAGCTTCGCACTCTCTTGGACTACTTGGAGCTCTAGAGAATTCCTCCGGGGGAAGATT 1320  
QY 1321 GTCCGCTCTGTGAGGGAGCCAGAGCACATCCCGACAAAGTTTCAGATAACCCCGTGG 1380  
DB 1321 GTCCGCTCTGTGAGGGAGCCAGAGCACATCCCGACAAAGTTTCAGATAACCCCGTGG 1380  
QY 1381 GATAGTTTTGCTAGATTTTGAAGAGGGGGGACAGAGGATGCTTCTTCAGTGCT 1440  
DB 1381 GATAGTTTTGCTAGATTTTGAAGAGGGGGGACAGAGGATGCTTCTTCAGTGCT 1440  
QY 1441 GAAATCCAGCGCAGAAATGTCGGTGTTCATGAAGTCACTGATTAGTTCCTGCTGCTGCTG 1500  
DB 1441 GAAATCCAGCGCAGAAATGTCGGTGTTCATGAAGTCACTGATTAGTTCCTGCTGCTGCTG 1500  
QY 1501 CAGGCAATAATAGGCGGCTCTCTTGGGACGACACTCTTCATACTAAGTGTGACATA 1560  
DB 1501 CAGGCAATAATAGGCGGCTCTCTTGGGACGACACTCTTCATACTAAGTGTGACATA 1560  
QY 1561 CACTGGATTTTGCATTTTTCAGGGTTCAGTATCCTATGACATGATGGTGTACCATCT 1620  
DB 1561 CACTGGATTTTGCATTTTTCAGGGTTCAGTATCCTATGACATGATGGTGTACCATCT 1620  
QY 1621 AATTTTACAGGCAAGACACAGCTGGAGAGTACAGAACTTGTCCCAAGGCTCAC 1680  
DB 1621 AATTTTACAGGCAAGACACAGCTGGAGAGTACAGAACTTGTCCCAAGGCTCAC 1680  
QY 1681 AGCCAGTAGGATAGAGCGGGAATGAAATCGAGCACTGTGAGAACTCTGGTGGGAGCC 1740  
DB 1681 AGCCAGTAGGATAGAGCGGGAATGAAATCGAGCACTGTGAGAACTCTGGTGGGAGCC 1740  
QY 1741 CCTGACTTGAACCACTCCCACTGCTCCCTTTCAGGAGGGACACTGATGATGAGTGC 1800  
DB 1741 CCTGACTTGAACCACTCCCACTGCTCCCTTTCAGGAGGGACACTGATGATGAGTGC 1800  
QY 1801 TCGAGCGGCGATCTTCCATCCCTGTTCAGTCCCTCCCTCCCTCCCTCCCTCCCTCCCTTGT 1860



Db 1801 TCGAGCCGCGCATCTTCCATCCCTGTCGAGTCCCTCCAGCTCCAGCTCCAGCTTGT 1860  
QY 1861 GCTTTTGGAGCTAAGCCTGGGATGACAAATTCACCCAGCTCCTTCAATCAGAGGCT 1920  
Db 1861 GCTTTTGGAGCTAAGCCTGGGATGACAAATTCACCCAGCTCCTTCAATCAGAGGCT 1920  
QY 1921 GGATGTAGCTGGGATGAGTCCATGTTATCGGCTCGGTACTCAACACAAACCAAGTTTCA 1980  
Db 1921 GGATGTAGCTGGGATGAGTCCATGTTATCGGCTCGGTACTCAACACAAACCAAGTTTCA 1980  
QY 1981 TCGAGGAAATGTCCCGCAGTGGATGAGCTCAGCTCAGCTGAGGAAACCCAGCTCTGA 2040  
Db 1981 TCGAGGAAATGTCCCGCAGTGGATGAGCTCAGCTCAGCTGAGGAAACCCAGCTCTGA 2040  
QY 2041 CAGAGTTCTTATAAATGTATAAATAGGCTCAGAAACCACTGCATTCGACCTGTGTAC 2100  
Db 2041 CAGAGTTCTTATAAATGTATAAATAGGCTCAGAAACCACTGCATTCGACCTGTGTAC 2100  
QY 2101 AGACTGCCACACATGCTGACCTGCCTAGCGAGGAGGACATCCCTTCTGAGCCATCTGCG 2160  
Db 2101 AGACTGCCACACATGCTGACCTGCCTAGCGAGGAGGACATCCCTTCTGAGCCATCTGCG 2160  
QY 2161 CTCTCTCATTTTCATCCCACTGCTCCCTGTTTGTGATCAATGGGACCACTGCTGC 2220  
Db 2161 CTCTCTCATTTTCATCCCACTGCTCCCTGTTTGTGATCAATGGGACCACTGCTGC 2220  
QY 2221 CCAGGAGCACATTTAGGGCTCTCAGTTCAAACTGAAGGACACAGTTGAACTCAGATGGGTT 2280  
Db 2221 CCAGGAGCACATTTAGGGCTCTCAGTTCAAACTGAAGGACACAGTTGAACTCAGATGGGTT 2280  
QY 2281 CATGTGGGATTCGGAGCTTCTGGGAATTCAGTTGGAGTCAAGTCAAGTCAAGTCAAGT 2340  
Db 2281 CATGTGGGATTCGGAGCTTCTGGGAATTCAGTTGGAGTCAAGTCAAGTCAAGTCAAGT 2340  
QY 2341 GGACCCCTCGGGCTCAGAGCCCTAAAGTGGCCCTGGTGAAGCAGGGTGGTCTCGGCTC 2400  
Db 2341 GGACCCCTCGGGCTCAGAGCCCTAAAGTGGCCCTGGTGAAGCAGGGTGGTCTCGGCTC 2400  
QY 2401 ACTTCCCAAGCCTGAGCCAAAGCTCATCTTCAATGAATGTCTCATTTGGCCGAGGAACAC 2460  
Db 2401 ACTTCCCAAGCCTGAGCCAAAGCTCATCTTCAATGAATGTCTCATTTGGCCGAGGAACAC 2460  
QY 2461 TGAACCTTTGGTGTGCTGTTTGGCTTGTGCTCCGCTGCTCCTACCCAGAGTT 2520  
Db 2461 TGAACCTTTGGTGTGCTGTTTGGCTTGTGCTCCGCTGCTCCTACCCAGAGTT 2520  
QY 2521 TGTGGAGCTGTGTTGAGGGTGTATAAACCAGGTACTTGTAGTTTGGCCATT 2580  
Db 2521 TGTGGAGCTGTGTTGAGGGTGTATAAACCAGGTACTTGTAGTTTGGCCATT 2580  
QY 2581 CAGCCATGCTCAGTGCATGCAAGTAACTTGTCTCTTAATATAGAAATGATTTTCT 2640  
Db 2581 CAGCCATGCTCAGTGCATGCAAGTAACTTGTCTCTTAATATAGAAATGATTTTCT 2640  
QY 2641 TTTAATTTTTTACTTTACCAGACTTTACTTTGTACTCAGAGAAGGCTCAGATGGCTG 2700  
Db 2641 TTTAATTTTTTACTTTACCAGACTTTACTTTGTACTCAGAGAAGGCTCAGATGGCTG 2700  
QY 2701 TGTACATATAAATGTGGACTAACTTTTAAAAAATAAATAAATAAATAAATAAATAA 2760  
Db 2701 TGTACATATAAATGTGGACTAACTTTTAAAAAATAAATAAATAAATAAATAAATAA 2760  
QY 2761 AAAAAAATAAATAA 2775  
Db 2761 AAAAAAATAAATAA 2775

RESULT 2  
AAA93876  
ID AAA93876 standard; DNA; 10562 BP.  
XX AC  
AC AC  
XX A.A93876;

DT 15-JAN-2001 (first entry)  
XX Human beta3Gal-T5 encoding DNA.  
DE Beta-1,3 galactose transferase; treatment; diagnosis; cancer; human;  
KW digestive system; beta3Gal-T5; ds.  
KW Homo sapiens.  
OS WO2000050608-A1.  
PN 31-AUG-2000.  
XX 24-FEB-2000; 2000WO-JP01070.  
XX 25-FEB-1999; 99JP-0047571.  
PR (KYOW ) KYOWA HAKKO KOGYO KK.  
XX Narimatsu H, Isshiki S, Togayachi A, Sasaki K;  
XX WPI; 2000-549409/50.  
DR Beta-1,3 galactose transferase and DNA encoding it, useful for  
PT synthesis of type 1 sialyl Lewis, a carbohydrate for treatment of  
PT digestive system cancer  
XX Claim 31; Page 103-111; 123pp; Japanese.  
XX This invention relates to a polypeptide (I) with beta-1,3 galactose  
CC transferase activity, or variants of (I) comprising amino acid additions,  
CC deletions and/or substitutions. Included in the invention is DNA encoding  
CC all or part of (I); expression vectors containing the DNA, host cells  
CC transformed by the vectors; a method for the preparation of the  
CC polypeptide by culture of the transformants or by expression in the milk  
CC of a transgenic mammal, and antibodies recognising (I). The Beta-1,3  
CC galactose transferase protein transfers galactose by beta-1,3 bonding to  
CC N-acetylglucosamine present in a non-cyclic carbohydrate chain (such as  
CC GlcNAc-beta1-3Gal-beta1-4Glc) to give Galbeta1-3GlcNAc. The protein and  
CC DNA encoding it are useful for the treatment and diagnosis of cancer of  
CC the digestive system. The present sequence represents a Beta3Gal-T5  
CC encoding DNA sequence.  
SQ Sequence 10562 BP; 2610 A; 2415 C; 2574 G; 2963 T; 0 other;  
Query Match 83.9%; Score 2329; DB 21; Length 10562;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 402 ATGGCTTTCCGAGATGAGATTGATGATATTTGCTTCTGGTCTGGGGGCTCTTTGT 461  
Db 8234 ATGGCTTTCCGAGATGAGATTGATGATATTTGCTTCTGGTCTGGGGGCTCTTTGT 8293  
QY 462 TTGTATTTAGCATGTACAGTCTAAATCCTTTCAAGAACAGTCTCTTTTACAGAAA 521  
Db 8294 TTGTATTTAGCATGTACAGTCTAAATCCTTTCAAGAACAGTCTCTTTTACAGAAA 8353  
QY 522 GACGGGAATTCCTTAAGCTCCAGATACAGACTGCAGGACAGACACTCCCTTCTCTGTC 581  
Db 8354 GACGGGAATTCCTTAAGCTCCAGATACAGACTGCAGGACAGACACTCCCTTCTCTGTC 8413  
QY 582 CTGCTGTGTACCTCATCCCAAAACAGTTGGCTGAGCGCATGGCCATCCGGCAGACGTGG 641  
Db 8414 CTGCTGTGTACCTCATCCCAAAACAGTTGGCTGAGCGCATGGCCATCCGGCAGACGTGG 8473  
QY 642 GGGAAAGAGAGGATGGTGAAGGAAAGCAGCTGAAGACATTTCTTCTCTGGGACACC 701  
Db 8474 GGGAAAGAGAGGATGGTGAAGGAAAGCAGCTGAAGACATTTCTTCTCTGGGACACC 8533  
QY 702 AGCAGTCCAGCGGAAACAGAGAGGTGACCGAGGAGACCGACGACGGGGACATTATC 761  
Db 8534 AGCAGTCCAGCGGAAACAGAGAGGTGACCGAGGAGACCGACGACGGGGACATTATC 8593



KW glycoprotein.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 79..1011  
FT /\*tag= a  
FT /product= "Beta3Gal-T5"  
FT primer\_bind complement (79..98)  
FT /\*tag= b  
FT primer\_bind complement (150..170)  
FT /\*tag= c  
FT primer\_bind 991..1011  
FT /\*tag= d  
XX  
PN WO200029558-A1.  
XX  
XX 25-MAY-2000.  
XX  
PF 11-NOV-1999; 99NO-US26807.  
XX  
XX 13-NOV-1998; 98DK-0001483.  
XX  
PA (CLAU/) CLAUSEN H.  
XX  
PI Clausen H, Amado M;  
XX  
XX WPI: 2000-399728/34.  
DR P-PSDB; AAY94641.  
XX  
PT Novel nucleic acid sequence encoding human  
PT UDP-galactose:beta-N-acetylglucosamine beta1,3-galactosyltransferase  
PT useful for obtaining beta 1,3-galactosyl glycosylated saccharides and  
PT glycopeptides or glycoproteins -  
XX  
PS Claim 7; Fig 1; 74pp; English.  
XX  
XX The present invention relates to a nucleic acid sequence encoding  
CC UDP-D-galactose:beta-N-acetylglucosamine beta1,3-galactosyltransferase  
CC (beta3Gal-T5). Beta3 transferases add galactose to the hydroxy group at  
CC carbon 3 of 2-acetamido-2-deoxy-D-glucose (GlcNAc). The present sequence  
CC represents the human beta3Gal-T5 gene sequence. The beta3Gal-T5 gene is  
CC located on human chromosome 21q22.3. Beta3Gal-T5 is a type II  
CC transmembrane glycoprotein. The invention also relates to the beta3Gal-T5  
CC protein sequence, a nucleic acid vector comprising the beta3Gal-T5  
CC nucleotide sequence, a host cell comprising the beta3Gal-T5  
CC the production of the beta3Gal-T5 protein from the host cells. The  
CC methods of the invention can be used for recombinant production of  
CC beta3Gal-T5 for use as a catalyst and for recombinant production of  
CC peptides or proteins with appropriate galactosylation. The beta3Gal-T5  
CC protein can be used to obtain beta1,3-galactosyl glycosylated  
CC saccharides, glycopeptides or glycoproteins.  
XX  
SQ Sequence 1011 BP; 247 A; 256 C; 251 G; 257 T; 0 other;

Query Match 28.8%; Score 798; DB 21; Length 1011;  
Best Local Similarity 99.8%; Pred. No. 5.3e-271;  
Matches 898; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 435 TGCCTCTCGTCTCGGGGCTCTTTGTTGTTATTTAGCATGTACAGTCTAAATCCTTTC 494  
Db 112 TGCCTCTCGTCTCGGGGCTCTTTGTTGTTATTTAGCATGTACAGTCTAAATCCTTTC 171  
QY 495 AAAGAACAGTCCCTTTGTTTACAAGAAAGACGGGAACCTTCCCTTAAGCTCCAGATACAGAC 554  
Db 172 AAAGAACAGTCCCTTTGTTTACAAGAAAGACGGGAACCTTCCCTTAAGCTCCAGATACAGAC 231  
QY 555 TGCAGGCACACACCTCCCTTCCTCGTCTCGTGTGACCTCATCCCAACAACAGTTGGCT 614  
Db 232 TGCAGGCACACACCTCCCTTCCTCGTCTCGTGTGACCTCATCCCAACAACAGTTGGCT 291  
QY 615 GAGCGCATGCCATCCGGCAGACGTGGGGAAAGAGAGGATGGTCAAGGGAAAGCAGCTG 674

Db 292 GAGCGCATGCCATCCGGCAGACAGCTGGGGGAAAGAGAGAGGAGCGGTGAAGGGGAAACGACGTG 351  
QY 675 AAGACATTCTTCTCTCTCTGGGGACCAACAGCAGTGCAGCGGAAACGAAAGAGGTGGACACG 734  
Db 352 AAGACATTCTTCTCTCTCTGGGGACCAACAGCAGTGCAGCGGAAACGAAAGAGGTGGACACG 411  
QY 735 GAGAGCCAGCAGACAGCGGGACATTTATCCAGAGGATTTCTAGACGTCTATTACAATCTG 794  
Db 412 GAGAGCCAGCAGACAGCGGGACATTTATCCAGAGGATTTCTAGACGTCTATTACAATCTG 471  
QY 795 ACCCTGAAGACCATGATGGGCATAGATGGTCCATCGCTTTTCTCTCAGCGCGCGTTT 854  
Db 472 ACCCTGAAGACCATGATGGGCATAGATGGTCCATCGCTTTTCTCTCAGCGCGCGTTT 531  
QY 855 GTGATGAAAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTTCTG 914  
Db 532 GTGATGAAAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTTCTG 591  
QY 915 AAGAAAAACAGAACCAACAGGTTTTTCACTGGCTTCTTGAACACTCAATGAGTTTCCCATC 974  
Db 592 AAGAAAAACAGAACCAACAGGTTTTTCACTGGCTTCTTGAACACTCAATGAGTTTCCCATC 651  
QY 975 AGGCAGCCATTTCAGCAAGTGGTTTGTCAAGTAATCTGAATATCGTGGGACAGGTACCCA 1034  
Db 652 AGGCAGCCATTTCAGCAAGTGGTTTGTCAAGTAATCTGAATATCGTGGGACAGGTACCCA 711  
QY 1035 CCATTCTGCTCGGCACCGCTACGTTTCTCTGGGACGTGGGAGTCAAGTGTACAAT 1094  
Db 712 CCATTCTGCTCGGCACCGCTACGTTTCTCTGGGACGTGGGAGTCAAGTGTACAAT 771  
QY 1095 GTCTCCAAAGAGGTCCTACATACATTAAACTGGAAGACGTGTTTGTGGGGCTCTGCCTCGAA 1154  
Db 772 GTCTCCAAAGAGGTCCTACATACATTAAACTGGAAGACGTGTTTGTGGGGCTCTGCCTCGAA 831  
QY 1155 AGGCTGAAACATCAGATTGGAGGAGCTCCACTCCCAGCCGACCTTTTTCAGGGGGGCTTA 1214  
Db 832 AGGCTGAAACATCAGATTGGAGGAGCTCCACTCCCAGCCGACCTTTTTCAGGGGGGCTTA 891  
QY 1215 CGCTTCTCCGTATGCTCTTTCAGGAGGATGCTGGCCCTGCCACTTCATCAAGCCCTCGGACT 1274  
Db 892 CGCTTCTCCGTATGCTCTTTCAGGAGGATGCTGGCCCTGCCACTTCATCAAGCCCTCGGACT 951  
QY 1275 CTCCTGGACTACTGCGAGGCTCTAGAGAATTCCCGGGGGGAAAGATTGTCGCGCTGTCTGA 1334  
Db 952 CTCCTGGACTACTGCGAGGCTCTAGAGAATTCCCGGGGGGAAAGATTGTCGCGCTGTCTGA 1011

RESULT 4  
ABN47639  
ID ABN47639 standard; DNA; 60 BP.  
XX AC  
XX ABN47639;  
XX AC  
XX DT 15-JUL-2002 (first entry)  
XX DE  
XX Human spliced transcript detection oligonucleotide SEQ ID NO:20387.  
XX KW  
XX Human; mouse; rat; splice transcript; detection; RNA transcript;  
XX splice variant; transcriptome; oligonucleotide library; ss.  
XX OS  
XX Homo sapiens.  
XX PN  
XX WO200210449-A2.  
XX PD  
XX 07-FEB-2002.  
XX PF  
XX 20-JUL-2001; 2001WO-IB01903.  
XX PR  
XX 28-JUL-2000; 2000US-221607P.  
XX PR  
XX 02-MAY-2001; 2001US-287724P.  
XX PA  
XX (COMP-) COMPUGEN INC.  
XX

```
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX WPI; 2002-257383/30.
XX
XX New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes
XX
XX Example 1; SEQ ID 20387; 47pp; English.
XX
XX The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN95589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 60 BP; 13 A; 8 C; 21 G; 18 T; 0 other;
SQ
Query Match 2.2%; Score 60; DB 24; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.3e-11;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 2265 GAATCAGATGGGTTTCATGGGATTCGGAGCTTCTGGGAATTCAGTTGGATCAA 2324
Db 1 GAATCAGATGGGTTTCATGGGATTCGGAGCTTCTGGGAATTCAGTTGGATCAA 60
RESULT 5
AAQ99552
XX AAQ99552 standard; cDNA to mRNA; 861 BP.
XX AC AAQ99552;
XX
XX 29-APR-1996 (first entry)
XX
XX Thrombopoietin coding sequence.
XX
XX Thrombopoietin; TPO; humoral factor; platelet; antibody; therapy; AIDS;
KW thrombocytopenia; hypoplastic anaemia; thrombotic thrombocytopenia;
KW disseminated intravascular coagulation syndrome; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 25..783
XX /tag= a
XX /product= thrombopoietin
XX sig_peptide 25..87
XX /tag= b
XX mat_peptide 88..783
XX /tag= c
XX
XX W09521919-A2.
XX
XX
```

```
XX 17-AUG-1995.
PD
XX 14-FEB-1995; 95WO-JP00208.
XX
XX 14-FEB-1994; 94JP-0039090.
PR 14-MAR-1994; 94US-0212164.
PR 25-MAR-1994; 94JP-0079842.
PR 01-APR-1994; 94US-0221020.
PR 15-JUN-1994; 94JP-0155126.
PR 20-JUL-1994; 94US-0278083.
PR 17-AUG-1994; 94JP-0193169.
PR 18-AUG-1994; 94JP-0193916.
PR 11-OCT-1994; 94US-0320300.
PR 01-NOV-1994; 94JP-0304167.
PR 01-DEC-1994; 94JP-0298669.
PR 22-DEC-1994; 94US-0361811.
PR 28-DEC-1994; 94JP-0341200.
PR 31-JAN-1995; 95US-0381478.
PR 17-AUG-1994; 94JP-0227159.
XX
XX (KIRI ) KIRIN BREWERY KK.
XX
XX Akahori H, Iwamatsu A, Kato T, Kuroki R, Miyazaki H;
PI Muto T, Ohgami K, Shimizu T;
XX
XX WPI; 1995-293120/38.
DR P-PSDB; AAR81378.
XX
XX Thrombopoietin polypeptide which specifically stimulates or
PT increases platelet production - useful in treatment of
PT thrombocytopenia, also related DNA and vectors
XX
XX Claim 19; Page 337-338; 383pp; English.
XX
XX This sequence represents the coding sequence for a thrombopoietin (TPO)
CC protein. TPO is a humoral factor capable of promoting platelet
CC production. The encoded protein can be purified by using a gel affinity
CC column where Mpl has been coupled to a resin. AAQ99551 and AAQ99553
CC also encode TPO proteins. The DNA sequences are inserted into vectors
CC which are used to transform prokaryotic and eukaryotic host cells. Using
CC the proteins encoded by these sequences, and derivatives of them,
CC antibodies specifically immunoreactive with a TPO protein can be created.
CC The antibodies, DNA sequences and vectors are used to isolate the
CC protein sequences. The TPO proteins can then be used in the treatment
CC of platelet disorders. These include thrombocytopenia, hypoplastic
CC anaemia, AIDS, disseminated intravascular coagulation syndrome and
CC thrombotic thrombocytopenia.
XX
XX Sequence 861 BP; 243 A; 257 C; 197 G; 164 T; 0 other;
SQ
Query Match 1.9%; Score 52; DB 16; Length 861;
Best Local Similarity 100.0%; Pred. No. 4.7e-09;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 2724 AACTCTTAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2775
Db 779 AACTCTTAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 830
RESULT 6
ABV54702
XX ABV54702 standard; cDNA; 190 BP.
XX
XX AC ABV54702;
XX
XX 17-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 54693.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
KW
```

```
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US05171.
XX PR 17-FEB-2000; 2000US-183319P.
XX PR 16-MAR-2000; 2000US-189862P.
XX PR 25-MAY-2000; 2000US-207454P.
XX PR 09-JUN-2000; 2000US-211314P.
XX PR 18-JUL-2000; 2000US-219007P.
XX PR 13-DEC-2000; 2000US-255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX DR WPI; 2001-662795/76.
XX PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX PS Claim 1; Page 10572; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX SQ Sequence 190 BP; 109 A; 28 C; 20 G; 33 T; 0 other;

Query Match 1.8%; Score 51; DB 23; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2725 ACTCTTAAAAA... 2775
DB 122 ACTCTTAAAAA... 172

RESULT 7
ABV55403
ID ABV55403 standard; cDNA; 390 BP.
XX AC ABV55403;
XX KW Human prostate expression marker cDNA 55394.
XX DT 17-SEP-2002 (first entry)
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
```

```
XX PF 20-FEB-2001; 2001WO-US05171.
XX PR 17-FEB-2000; 2000US-183319P.
XX PR 16-MAR-2000; 2000US-189862P.
XX PR 25-MAY-2000; 2000US-207454P.
XX PR 09-JUN-2000; 2000US-211314P.
XX PR 18-JUL-2000; 2000US-219007P.
XX PR 13-DEC-2000; 2000US-255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX DR WPI; 2001-662795/76.
XX PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX PS Claim 1; Page 10700; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX SQ Sequence 390 BP; 162 A; 83 C; 62 G; 83 T; 0 other;

Query Match 1.8%; Score 51; DB 23; Length 390;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2725 ACTCTTAAAAA... 2775
DB 180 ACTCTTAAAAA... 230

RESULT 8
ABV54507
ID ABV54507 standard; cDNA; 422 BP.
XX AC ABV54507;
XX DT 17-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 54498.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US05171.
XX PR 17-FEB-2000; 2000US-183319P.
XX PR 16-MAR-2000; 2000US-189862P.
XX PR 25-MAY-2000; 2000US-207454P.
```

PR 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
PR 13-DEC-2000; 2000US-255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
XX  
PI Schlegel R, Endege WO, Monahan JE;  
XX  
XX WPI; 2001-662795/76.  
XX  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer -  
XX  
PS Claim 1; Page 10535; 11750pp; English.  
XX  
XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 422 BP; 176 A; 92 C; 63 G; 91 T; 0 other;  
Query Match 1.8%; Score 51; DB 23; Length 422;  
Best Local Similarity 100.0%; Pred. No. 1.2e-08;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2725 ACTCTTAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2775  
Db 181 ACTCTTAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 231  
RESULT 9  
ABV19654  
ID ABV19654 standard; cDNA; 215 BP.  
XX  
AC ABV19654;  
XX  
DT 13-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 19645.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US05171.  
XX  
XX 17-FEB-2000; 2000US-183319P.  
PR 16-MAR-2000; 2000US-189862P.  
PR 25-MAY-2000; 2000US-207454P.  
PR 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
PR 13-DEC-2000; 2000US-255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
XX

PI Schlegel R, Endege WO, Monahan JE;  
XX  
XX WPI; 2001-662795/76.  
XX  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer -  
XX  
PS Claim 1; Page 3211-3212; 11750pp; English.  
XX  
XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 215 BP; 115 A; 17 C; 18 G; 57 T; 8 other;  
Query Match 1.8%; Score 50; DB 23; Length 215;  
Best Local Similarity 100.0%; Pred. No. 3.2e-08;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2726 CTCTTAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2775  
Db 142 CTCTTAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 191  
RESULT 10  
ABV49421  
ID ABV49421 standard; cDNA; 341 BP.  
XX  
AC ABV49421;  
XX  
DT 17-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 49412.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US05171.  
XX  
XX 17-FEB-2000; 2000US-183319P.  
PR 16-MAR-2000; 2000US-189862P.  
PR 25-MAY-2000; 2000US-207454P.  
PR 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
PR 13-DEC-2000; 2000US-255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
XX Schlegel R, Endege WO, Monahan JE;  
XX  
XX WPI; 2001-662795/76.  
XX  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer, useful

PT for detecting presence of prostate cancer, stage of prostate cancer -  
XX Claim 1; Page 9653; 11750pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 341 BP; 144 A; 64 C; 54 G; 79 T; 0 other;

Query Match 1.8%; Score 50; DB 23; Length 341;  
Best Local Similarity 100.0%; Pred. No. 2.9e-08;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2726 CTCCTTAAAAA... 2775  
DB 180 CTCCTTAAAAA... 229

RESULT 11  
AAI88530  
ID AAI88530 standard; cDNA; 383 BP.  
XX  
AC AAI88530;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 8590.  
XX  
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200164835-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 26-FEB-2001; 2001WO-US04927.  
XX  
PR 28-FEB-2000; 2000US-0515126.  
PR 18-MAY-2000; 2000US-0577409.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
XX WPI; 2001-514838/56.  
DR P-PSDB; AAO08599.  
XX  
XX Isolated nucleic acids and polypeptides, useful for preventing  
PT diagnosing and treating e.g. leukaemia, inflammation and immune  
PT disorders -  
XX  
XX Claim 1; SEQ ID NO 8590; 1399pp + Sequence Listing; English.  
XX  
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.

CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 383 BP; 192 A; 74 C; 65 G; 52 T; 0 other;

Query Match 1.8%; Score 50; DB 22; Length 383;  
Best Local Similarity 100.0%; Pred. No. 2.8e-08;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2726 CTCCTTAAAAA... 2775  
DB 106 CTCCTTAAAAA... 155

RESULT 12  
AAI83204  
ID AAI83204 standard; cDNA; 386 BP.  
XX  
AC AAI83204;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 3264.  
XX  
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200164835-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 26-FEB-2001; 2001WO-US04927.  
XX  
PR 28-FEB-2000; 2000US-0515126.  
PR 18-MAY-2000; 2000US-0577409.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
XX WPI; 2001-514838/56.  
DR P-PSDB; AAO03273.  
XX  
XX Isolated nucleic acids and polypeptides, useful for preventing  
PT diagnosing and treating e.g. leukaemia, inflammation and immune  
PT disorders -  
XX  
XX Claim 1; SEQ ID NO 3264; 1399pp + Sequence Listing; English.  
XX  
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.

## RESULT 15



ABV57740  
ID ABV57740 standard; cDNA; 430 BP.  
XX  
AC ABV57740;  
XX  
DT 17-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 57731.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US05171.  
XX  
PR 17-FEB-2000; 2000US-183319P.  
PR 16-MAR-2000; 2000US-189862P.  
PR 25-MAY-2000; 2000US-207454P.  
PR 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
PR 13-DEC-2000; 2000US-255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JE;  
XX  
DR WPI: 2001-662795/76.  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer -  
XX  
PS Claim 1: Page 11099-11100; 11750pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 430 BP; 196 A; 89 C; 85 G; 60 T; 0 other;

Query Match 1.8%; Score 50; DB 23; Length 430;  
Best Local Similarity 100.0%; Pred. No. 2.8e-08;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2726 CTCCTTAAAAA... 2775  
Dd 124 CTCCTTAAAAA... 173

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 12, 2003, 02:50:35 ; Search time 3046.33 Seconds  
(without alignments)  
14752.993 Million cell updates/sec

Title: US-09-914-152-2  
Perfect score: 2775  
Sequence: 1 gtgaattccttctctgc.....aaaaaaaaaaaaaaaaaaaaa 2775

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST.\*

- 1: em\_estba.\*
- 2: em\_esthum.\*
- 3: em\_estin.\*
- 4: em\_estmu.\*
- 5: em\_estov.\*
- 6: em\_estpl.\*
- 7: em\_estro.\*
- 8: em\_htc.\*
- 9: gb\_est1.\*
- 10: gb\_est2.\*
- 11: gb\_htc.\*
- 12: gb\_est3.\*
- 13: gb\_est4.\*
- 14: gb\_est5.\*
- 15: em\_estfun.\*
- 16: em\_estom.\*
- 17: gb\_gss.\*
- 18: em\_gss\_hum.\*
- 19: em\_gss\_inv.\*
- 20: em\_gss\_pln.\*
- 21: em\_gss\_vrt.\*
- 22: em\_gss\_fun.\*
- 23: em\_gss\_mam.\*
- 24: em\_gss\_mus.\*
- 25: em\_gss\_other.\*
- 26: em\_gss\_pro.\*
- 27: em\_gss\_rod.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	461	16.6	507	9	AI739518
C 2	455	16.4	455	10	AW136254
C 3	215	7.7	434	12	BF748181
C 4	181	6.5	528	17	AQ266346
C 5	165	5.9	338	9	AJ003597
C 6	104	3.7	108	12	BE699666

C91526	C91526	Dict
BE047732	tz41d08.y	
B1872028	603396322	
BM920507	AGENCOURT	
AL546967	AL546967	
BG632911	GH15835.5	
BM154294	fv83d01.y	
AA761919	nz42h05.s	
BE198722	u981g09.y	
W36406	mb83f10.r1	
BC007903	Homo_sapi	
AA478826	21787.WAR	
BF945160	SB29d04.y	
AW344775	26507.MAR	
BE012908	122634.MA	
BE236290	144016.MA	
BF191579	239203.MA	
B1823175	603039425	
AW785638	116976.MA	
BG382634	298603.MA	
BG895631	359178.MA	
AV759831	AV759831	
BE233664	140100.MA	
BG383966	302440.MA	
B1945683	SC47a07.y	
BG834738	353063.MA	
BG754013	602709541	
BG717868	AGENCOURT	
AW325697	17623.MAR	
BF811802	MR2-C1018	
BG638107	hdl18e10.y	
AI590535	tw11b06.x	
AW432772	sh82g05.y	
AL040602	DKFp434L	
BG153044	nah25b12.	
BG981758	MR3-CN014	
BE666981	150939.MA	
BG058273	nah23a02.	
BF749517	IL0-BN042	

ALIGNMENTS

RESULT 1	AI739518/c	
LOCUS	w123a06.x1 NCI_CGAP_Col6 Homo sapiens cDNA clone IMAGE:2391058 3', mRNA sequence.	507 bp
DEFINITION	AI739518	
ACCESSION	AI739518.1	GI:5101499
VERSION	EST.	
KEYWORDS	human.	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 507)	
AUTHORS	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .	
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index	
JOURNAL	Unpublished (1997)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://www.bio.llnl.gov/dbbrp/image/image.html">www.bio.llnl.gov/dbbrp/image/image.html</a> Insert Length: 1331 Std Error: 0.00 Seq primer: -40UP from Gibco	

High quality sequence stop: 465.  
Location/Qualifiers  
1. 507  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2391058"  
/clone\_lib="NCI\_CGAP\_Col6"  
/tissue\_type="colon tumor, RER-"  
/lab\_host="DH10B"  
/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not 1; Site\_2: Eco RI; Plasmid DNA from the normalized library NCI\_CGAP\_Col0 was prepared, and ss circles were made in vitro. Following HAP hybridization reaction, this DNA was used as a tracer in a subtractive library. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneDs 1057416-1061255, and 1144584-1145351). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 154 a 122 c 115 g 116 t  
ORIGIN

Query Match 16.6%; Score 461; DB 9; Length 507;  
Best Local Similarity 100.0%; Pred. No. 1.4e-94;  
Matches 461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2223 CAGGAGACATTTAGGGCTCTCAGTTCAACACGAGGAGGAGTTGAAGTCTCAGATGGGGTTCA 2282  
|||||  
Db 507 CAGGAGACATTTAGGGCTCTCAGTTCAACACGAGGAGGAGTTGAAGTCTCAGATGGGGTTCA 448  
|||||

QY 2283 TGTGGGATTCGGGAGCTTTCTGGGAATTCAGTTGAGTCAAGTCAAGGATGCTCTCAAGG 2342  
|||||  
Db 447 TGTGGGATTCGGGAGCTTTCTGGGAATTCAGTTGAGTCAAGTCAAGGATGCTCTCAAGG 388  
|||||

QY 2343 ACCCTCGGGCTCAGAGCCCTAAAGTGGGCGCTGTGGAAGCAGGGTGTGCTCGCTGCCAC 2402  
|||||  
Db 387 ACCCTCGGGCTCAGAGCCCTAAAGTGGGCGCTGTGGAAGCAGGGTGTGCTCGCTGCCAC 328  
|||||

QY 2403 TTCCCAAGCTGAGCCAGCTCCTTCATTAAGTGTCTCATTTGGCCGAGGAACAAGTG 2462  
|||||  
Db 327 TTCCCAAGCTGAGCCAGCTCCTTCATTAAGTGTCTCATTTGGCCGAGGAACAAGTG 268  
|||||

QY 2463 AACTTGTGTTGCTGCTTTAGCCCTCAGTTTGTGCTCCGCTGCTCTACCCAGAGGTTTG 2522  
|||||  
Db 267 AACTTGTGTTGCTGCTTTAGCCCTCAGTTTGTGCTCCGCTGCTCTACCCAGAGGTTTG 208  
|||||

QY 2523 TGGGAGCTGTGTGAGGCTGTATATAACCAAGTACTTCGTTAGTTTGTGCCATTTCA 2582  
|||||  
Db 207 TGGGAGCTGTGTGAGGCTGTATATAACCAAGTACTTCGTTAGTTTGTGCCATTTCA 148  
|||||

QY 2583 GCCATGCTACGTGACATGCAAGTAATCTTGTCTCTAATTATAGAAATGATTTTCTTT 2642  
|||||  
Db 147 GCCATGCTACGTGACATGCAAGTAATCTTGTCTCTAATTATAGAAATGATTTTCTTT 88  
|||||

QY 2643 TAATTTTACTTTACAGACTTTTACTTTGTACTCAGAGAA 2683  
|||||  
Db 87 TAATTTTACTTTACAGACTTTTACTTTGTACTCAGAGAA 47  
|||||

RESULT 2  
AW136254/c  
LOCUS  
DEFINITION UI-H-B11-act-e-12-0-UI.s1 NCI\_CGAP\_Sub3 Homo sapiens cDNA clone  
IMAGE:2715383 3', mRNA sequence.  
ACCESSION AW136254  
VERSION  
KEYWORDS EST.  
SOURCE  
ORGANISM human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 455)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP).

Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov  
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LiLi at: www.bio.lnlnl.gov/bbrp/image/image.html  
Seq primer: M13 Forward  
POLYA-yes.

Location/Qualifiers  
1. 455  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2715383"  
/clone\_lib="NCI\_CGAP\_Sub3"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not 1; Site\_2: Eco RI; The NCI\_CGAP\_Sub3 library is a subtracted library derived from the NCI\_CGAP\_Sub1 library, which is a subtracted library derived from B1. B1 constitutes a mixture of 21 normalized or subtracted NCI\_CGAP libraries: NCI\_CGAP\_Col0, NCI\_CGAP\_Col16, NCI\_CGAP\_Kid5, NCI\_CGAP\_Kid12, NCI\_CGAP\_Kid3, NCI\_CGAP\_Kid11, NCI\_CGAP\_Lym2, NCI\_CGAP\_Br2, NCI\_CGAP\_Col8, NCI\_CGAP\_CLL1, NCI\_CGAP\_Lei2, NCI\_CGAP\_Brn23, NCI\_CGAP\_Lu5, NCI\_CGAP\_Lu24, NCI\_CGAP\_Lu19, NCI\_CGAP\_G4, NCI\_CGAP\_G6, NCI\_CGAP\_Brn25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below:  
NCI\_CGAP\_Kid3 pool 1 LLAM 3334-3337, 3682-3683,  
3798-3803 (IMAGE Cloneds 132376-1323911, 1456008-1456775  
, 1500552-1502855); NCI\_CGAP\_Kid5 pool 1 LLAM 3338-3342  
, 3722-3725, 3776-3778 (IMAGE Cloneds 1323912-1325831,  
1471368-1472903, 1492104-1493255); NCI\_CGAP\_Lu5 pool 1,  
LLAM 3575-3582, 3851-3854 (IMAGE Cloneds 1414920-1417991,  
1520904-1522439); NCI\_CGAP\_G4 pool 1 LLAM 3164-3167,  
3716-3720, 3733-3735 (IMAGE Cloneds 1257096-1258631,  
1469064-1470983, 1475592-1476743); NCI\_CGAP\_Pr22 pool 1  
LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE Cloneds  
985608-986759, 1101192-1101959, 1217928-1220615);  
NCI\_CGAP\_Col10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE  
Cloneds 1057416-1061255, 1144584-1145351). Subtraction  
was performed as previously described (Bonaldo, Lennon &  
Soares (1996): Normalization and Subtraction: Two  
Approaches To Facilitate Gene Discovery. Genome Research  
6, 791-806.  
TAG\_LiLi-NCI\_CGAP\_Col0  
TAG\_TISSUE=colon  
TAG\_SEQ=AAACG"

BASE COUNT 137 a 102 c 101 g 115 t  
ORIGIN

Query Match 16.4%; Score 455; DB 10; Length 455;  
Best Local Similarity 100.0%; Pred. No. 3.3e-93;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2293 TGGGAGCTTTCTGGGAATTCAGTTGGAGTCAAGTCAAGGATGCTCTCAAGGACCCCTCGG 2352  
|||||  
Db 455 TGGGAGCTTTCTGGGAATTCAGTTGGAGTCAAGTCAAGGATGCTCTCAAGGACCCCTCGG 396  
|||||

QY 2353 CTCAGAGCCCTAAAGTGGCCCTGGTGAAGCAGGGTGGTCTCGGTCCACTTCCCAAGCC 2412  
|||||  
Db 395 CTCAGAGCCCTAAAGTGGCCCTGGTGAAGCAGGGTGGTCTCGGTCCACTTCCCAAGCC 336  
|||||

QY 2413 TGAGCCAGCTCATCTTCTATTCAATGCTCTCAATTGGCCGAGGACACTCAACTTTGGG 2472  
|||||

```
Db 335 TGAGCCAGCTCATCTTCATTGAATGCTCTCATTTGGCCGAGGAACAACCTGAACCTTTGTGG 276
Qy 2473 TTTGCTGTTTACCTTCAGTTTGCTCCGCTGCTCTACCCAGAGGTTTGTGCGAGCGTG 2532
Db 275 TTTGCTGTTTACCTTCAGTTTGCTCCGCTGCTCTACCCAGAGGTTTGTGCGAGCGTG 216
Qy 2533 TGTTCAGGTTGTATATAAACAAGTACTTCGTTAGTTTGGCCCATTCAGCCATGTGCA 2592
Db 215 TGTTCAGGTTGTATATAAACAAGTACTTCGTTAGTTTGGCCCATTCAGCCATGTGCA 156
Qy 2593 CGTGACATGCAAGTAATCTTCTCTCTAATATAGAAATGATTTCTTTTAAATTTTAA 2652
Db 155 CGTGACATGCAAGTAATCTTCTCTCTAATATAGAAATGATTTCTTTTAAATTTTAA 96
Qy 2653 CTTTACCAGACTTTACTTTGTACTCAGAGAAGGCGCTCACATGGCTGTGTGCACATATA 2712
Db 95 CTTTACCAGACTTTACTTTGTACTCAGAGAAGGCGCTCACATGGCTGTGTGCACATATA 36
Qy 2713 ATGTTGGACTAACTCTTAAAAAATAAAAAA 2747
Db 35 ATGTTGGACTAACTCTTAAAAAATAAAAAA 1

RESULT 3
BF748181/c
LOCUS
DEFINITION RC3-BN0411-021000-021-c04 BN0411 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF748181
VERSION BF748181.1 GI:12074857
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 434)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Ngai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
Goldman,G.H., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
Brunstein,A., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC3&t2=RC3-BN0411-
021000-021-c04&t3=2000-10-02&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 24
High quality sequence stop: 432.
Location/Qualifiers
1..434
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BN0411"
/dev_stage="Adult"
/note="Organ: breast normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
```

```
BASE COUNT 104 a 117 c 130 g 83 t
ORIGIN
Query Match 7.7%; Score 215; DB 12; Length 434;
Best Local Similarity 99.4%; Pred. No. 2.3e-39;
Matches 315; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 959 CAATGAGTTTCCCATCAGGAGGAGCCATTTCAGCAAGTGGTTTGTCAATAATCTGAATATCC 1018
Db 434 CAATGAGTTTCCCATCAGGAGGAGCCATTTCAGCAAGTGGTTTGTCAATAATCTGAATATCC 375
Qy 1019 GTGGGACAGTACCCACCATTCTCTCGGACCGGCTACGTGTTTCTTGGGACGCTGGC 1078
Db 374 GTGGGACAGTACCCACCATTCTCTCGGACCGGCTACGTGTTTCTTGGGACGCTGGC 315
Qy 1079 GAGTCAGGTGTACAATGCTTCCAGAGCGTCCCATACATTAAACTGGAAGAGCGTGTGT 1138
Db 314 GAGTCAGGTGTACAATGCTTCCAGAGCGTCCCATACATTAAACTGGAAGAGCGTGTGT 255
Qy 1139 GGGGCTCTGCTCGAAAGGCTGAACATCAGATTGGAGGAGCTCCACTCCCGACCGACCTT 1198
Db 254 GGGGCTCTGCTCGAAAGGCTGAACATCAGATTGGAGGAGCTCCACTCCCGACCGACCTT 195
Qy 1199 TTTTCCAGGGGCTTAGCTTCTCCGTATGCTCTTCCAGGAGGATCGTGGCTGCCACTT 1258
Db 194 TTTTCCGGGGGCTTAGCTTCTCCGTATGCTCTTCCAGGAGGATCGTGGCTGCCACTT 135
Qy 1259 CATCAAGCCTCGGACTC 1275
Db 134 CATCAAGCCTCGGACTC 118

RESULT 4
AQ266346
LOCUS
DEFINITION AQ266346 528 bp DNA linear GSS 27-APR-1999
RP0111-72B14.TJ RPI-11 Homo sapiens genomic clone RPI-11-72B14,
DNA sequence.
ACCESSION AQ266346
VERSION AQ266346.1 GI:3793950
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 528)
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
Use of human BAC End Sequences for Sequence-Ready Map Building
Unpublished (1998)
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library RPI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1..528
/organism="Homo sapiens"
/db_xref="GDB:7527301"
/db_xref="taxon:9606"
/clone_lib="RPI-11-72B14"
/clone_lib="RPI-11"
/sex="Male"
/cell_type="Lymphocytes"
```

/note="Vector: pBACE3.6; Site.1: EcoRI; Site.2: EcoRI;  
RPcIII Human Male BAC Library"  
BASE COUNT 101 a 146 c 104 g 175 t 2 others  
ORIGIN

Query Match 6.5%; Score 181; DB 17; Length 528;  
Best Local Similarity 100.0%; Pred. No. 8e-32;  
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 GAGCTGAGGCTCTCTGGCCCTTTGGACCCGAGGATTTATACCAAGCAGGTTTCTGGGTTT 120  
DB 56 GAGCTGAGGCTCTCTGGCCCTTTGGACCCGAGGATTTATACCAAGCAGGTTTCTGGGTTT 115

QY 121 TCAGGCCCTTGGCCCTTGAGCTAGTAGTTACACCATGTCATCTGGTCTGAGGCTCTT 180  
DB 116 TCAGGCCCTTGGCCCTTGAGCTAGTAGTTACACCATGTCATCTGGTCTGAGGCTCTT 175

QY 181 GGTCTTGAGCTGAGCCACACCTCCCTGGCATCCAGGCTTCCAGCTTGCAATGGCTCTGAC 240  
DB 176 GGTCTTGAGCTGAGCCACACCTCCCTGGCATCCAGGCTTCCAGCTTGCAATGGCTCTGAC 235

QY 241 G 241  
DB 236 G 236

RESULT 5  
AJ003597/c  
LOCUS  
DEFINITION AJ003597 Selected chromosome 21 cDNA library Homo sapiens cDNA  
clone MP19-1219, mRNA sequence.  
ACCESSION AJ003597  
VERSION AJ003597.1 GI:2578270  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 338)  
AUTHORS Szulzewsky,I., Hunt,E., Nguyen,M., Korn,B., Roehrdanz,B., Lehrach  
H. and Vaspo,M.L.H.  
TITLE An integrated transcript map for the whole human chromosome 21  
JOURNAL Unpublished (1997)  
COMMENT Contact: Yaspo, M.-L.  
Max Planck Institut fuer Molekulare Genetik  
Inhnestrasse 73, D14195 Berlin-Dahlem, Germany.  
FEATURES  
source  
1..338  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="21q"  
/clone="MP19-1219"  
/clone\_lib="Selected chromosome 21 cDNA library"  
/note="Swaroop et al. (1991) Nucleic Acids Res. 19: 1954."  
BASE COUNT 76 a 92 c 82 g 87 t 1 others  
ORIGIN

Query Match 5.9%; Score 165; DB 9; Length 338;  
Best Local Similarity 99.5%; Pred. No. 4.8e-28;  
Matches 215; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 439 TTCTGGTCTGGGGGCTCTTGTGTTGATTTTAGCATGTACAGTCTAAATCCTTTTCAAAG 498  
DB 312 TTCTGGTCTGGGGGCTCTTGTGTTGATTTTAGCATGTACAGTCTAAATCCTTTTCAAAG 253

QY 499 AACAGTCTTTGTTTACAAGAAAGACGGGAACCTTCTTAAAGTCCCGATACAGACTGCA 558  
DB 252 AACAGTCTTTGTTTACAAGAAAGACGGGAACCTTCTTAAAGTCCCGATACAGACTGCA 193

QY 559 GGCAGACACCTCCCTTCTCTGCTGCTGGTGAGCTATCCCAACAAACAGTTGGCTGAGC 618  
DB 192 GGCAGACACCTCCCTTCTCTGCTGCTGGTGAGCTATCCCAACAAACAGTTGGCTGAGC 133

QY 619 GCATGGCCATCCGGCAGAGCTGGGGGAAAGAGAGGA 654  
DB 132 GCATGGCCATCCGGCAGAGCTGGGGGAAAGAGAGGA 97

RESULT 6  
BE699666/c  
LOCUS  
DEFINITION BE699666 108 bp mRNA linear EST 12-SEP-2000  
CM4-NN0084-100600-195-f05 NN0084 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BE699666  
VERSION BE699666.1 GI:10087408  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 108)  
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovsky-Almeida,S., Briones,M.R.,  
Negai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,  
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.  
TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et3-CM4-NN0084-100  
600-195-f05&t3=2000-06-10&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 7  
High quality sequence stop: 108.  
Location/Qualifiers  
1..108  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NN0084"  
/dev\_stage="Adult"  
/note="Organ: nervous normal; Vector: puc18; Site.1: SmaI;  
Site.2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."  
BASE COUNT 21 a 35 c 27 g 25 t  
ORIGIN

Query Match 3.7%; Score 104; DB 12; Length 108;  
Best Local Similarity 100.0%; Pred. No. 7.2e-14;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1271 GACTCTCTTGAGCTACTGCGAGGCTCTAGAGAATCCCGGGGGAAGATTGTCCGCTGT 1330  
DB 108 GACTCTCTTGAGCTACTGCGAGGCTCTAGAGAATCCCGGGGGAAGATTGTCCGCTGT 49

QY 1331 CTGAGGGAGCCCGAGGCGACATCCGGGACAAAGTTTCAGATAACC 1374  
DB 48 CTGAGGGAGCCCGAGGCGACATCCGGGACAAAGTTTCAGATAACC 5

RESULT 7  
C91526

LOCUS C91526 171 bp mRNA linear EST 20-APR-1998  
DEFINITION C91526 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium  
discoideum cDNA clone SSK390, mRNA sequence.  
ACCESSION C91526  
VERSION C91526.1 GI:3060892  
KEYWORDS EST.  
SOURCE Dictyostelium discoideum.  
ORGANISM Dictyostelium discoideum.  
REFERENCE Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.  
AUTHORS Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,  
Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takenoto,K., Yasukawa,H.,  
Williams,J., Maeda,M., Takeuchi,I., Ochial,H. and Tanaka,Y.  
TITLE Developmental cDNA in Dictyostelium discoideum  
JOURNAL Unpublished (1998)  
COMMENT Contact: Hideko Urushihara  
Institute of Biological Sciences  
University of Tsukuba  
1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan  
Tel: 81-298-53-4664  
Fax: 81-298-53-6614  
Email: hideko@biol.tsukuba.ac.jp.  
FEATURES  
source  
1..171  
/organism="Dictyostelium discoideum"  
/strain="AX4"  
/db\_xref="taxon:44689"  
/clone="SSK390"  
/clone\_lib="Dictyostelium discoideum SS (H.Urushihara)"  
/dev\_stage="slug"  
BASE COUNT 106 a 15 c 1 g 49 t  
ORIGIN  
Query Match 2.0%; Score 55; DB 14; Length 171;  
Best Local Similarity 100.0%; Pred. No. 0.0044;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2721 CTAACCTCTTAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2775  
|||||  
Db 115 CTAACCTCTTAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 169  
RESULT 8  
LOCUS BE047732 267 bp mRNA linear EST 08-JUN-2000  
DEFINITION BE047732 t24id08.y1 NCI\_CGAP\_Brn52 Homo sapiens cDNA clone IMAGE:2291151 5',  
mRNA sequence.  
ACCESSION BE047732  
VERSION BE047732.1 GI:8364785  
KEYWORDS EST..  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 267)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov  
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael  
R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL, send email to:  
info@image.llnl.gov  
Seq primer: -40RP from Gibco.  
FEATURES  
source  
1..267  
/organism="Homo sapiens"

/db\_xref="taxon:9606"  
/clone="IMAGE:2291151"  
/clone\_lib="NCI\_CGAP\_Brn52"  
/tissue\_type="tumor, 5 pooled (see description)"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; This library represents the normalized  
version of NCI\_CGAP\_Brn35. Cloned unidirectionally.  
Primer: Oligo dt. Average insert size 1.19 kb. Tumor  
types include: meningioma, oligodendroglioma, astrocytoma  
(grade II), medulloblastoma, astrocytoma (grade IV).  
Constructed by Life Technologies."  
BASE COUNT 133 a 24 c 20 g 90 t  
ORIGIN  
Query Match 1.9%; Score 54; DB 10; Length 267;  
Best Local Similarity 100.0%; Pred. No. 0.0047;  
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2722 TAAACTCTTAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2775  
|||||  
Db 176 TAAACTCTTAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 229  
RESULT 9  
LOCUS BI872028 1281 bp mRNA linear EST 11-OCT-2001  
DEFINITION BI872028 60396322F1 NIH\_MGC\_94 Mus musculus cDNA clone IMAGE:5400316 5',  
mRNA sequence.  
ACCESSION BI872028  
VERSION BI872028.1 GI:16045703  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 1281)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov  
Tissue Procurement: The Cepko Laboratory  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM12020 row: k column: 05  
High quality sequence stop: 311.  
FEATURES  
source  
1..1281  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5400316"  
/clone\_lib="NIH\_MGC\_94"  
/tissue\_type="retina"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: eye; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally; oligo-dt primed.  
Average insert size 3.3 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."  
BASE COUNT 708 a 125 c 342 g 106 t  
ORIGIN  
Query Match 1.9%; Score 54; DB 13; Length 1281;  
Best Local Similarity 100.0%; Pred. No. 0.0009;  
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2722 TAAACTCTTAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2775  
|||||

```

Db      233 TAAACTCTTAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 306

RESULT 10
BM920507
LOCUS
DEFINITION  BM920507 326 bp mRNA linear EST 12-MAR-2002
AGENCOURT_6709647 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5750494
5', mRNA sequence.
ACCESSION  BM920507
VERSION    BM920507.1 GI:19370886
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 326)
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: c9apbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1AM12781 row: 1 column: 23
            High quality sequence stop: 327.
            Location/Qualifiers
                1..326
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:5750494"
                /clone_lib="NIH_MGC_122"
                /lab_host="DH10B"
                /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
                Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
                anonymous pool of 24 week female lung, 16 week female
                spleen, and 20-22 week male spleens. Library is oligo-dT
                primed and directionally cloned (EcoRV site is destroyed
                upon cloning). Average insert size 1.4 kb, insert size
                range 1-3 kb. Library is normalized and enriched for
                full-length clones and was constructed by C. Gruber
                (Invitrogen). Research Genetics tracking code 026. Note:
                this is a NIH_MGC Library."
                149 a 50 c 70 g 57 t

BASE COUNT  149 a 50 c 70 g 57 t
ORIGIN
Query Match 1.9%; Score 53; DB 14; Length 326;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2723 AAACCTCTTAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2775
|||||.....
Db 231 AAACCTCTTAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 283

RESULT 11
AL546967
LOCUS
DEFINITION  AL546967 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1032YN07 5
prime mRNA sequence.
ACCESSION  AL546967
VERSION    AL546967.1 GI:12880599
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 521)
            Li.W.B., Gruber,C., Jessee,J. and Polayes,D.

BASE COUNT  149 a 50 c 70 g 57 t
ORIGIN
Query Match 1.9%; Score 53; DB 14; Length 326;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2723 AAACCTCTTAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2775
|||||.....
Db 231 AAACCTCTTAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 283

RESULT 11
AL546967
LOCUS
DEFINITION  AL546967 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1032YN07 5
prime mRNA sequence.
ACCESSION  AL546967
VERSION    AL546967.1 GI:12880599
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 521)
            Li.W.B., Gruber,C., Jessee,J. and Polayes,D.

BASE COUNT  149 a 50 c 70 g 57 t
ORIGIN
Query Match 1.9%; Score 53; DB 14; Length 326;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2723 AAACCTCTTAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2775
|||||.....
Db 231 AAACCTCTTAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 283

RESULT 12
BG632911
LOCUS
DEFINITION  BG632911 208 bp mRNA linear EST 23-APR-2001
melanogaster cDNA clone GH15835 5, mRNA sequence.
ACCESSION  BG632911
VERSION    BG632911.1 GI:13758395
KEYWORDS   EST.
SOURCE     fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
            1 (bases 1 to 208)
            Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
            BDGP/HMI Drosophila EST Project
            Unpublished (2001)
            Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: GH.158 row: C column: 11
            High quality sequence stop: 206.
            Location/Qualifiers
                1..208
                /organism="Drosophila melanogaster"
                /db_xref="taxon:7227"
                /clone="GH15835"
                /clone_lib="GH Drosophila melanogaster head pot2"
                /sex="male and female"
                /dev_stage="adult"
                /lab_host="DH5-alpha"
                /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                XhoI; Sized fractionated cDNAs were directly ligated into
                pOT2. Plasmid cDNA library."

Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1..521
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1032YN07"
/clone_lib="LTI_NFL006.PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT  225 a 85 c 91 g 119 t 1 others
ORIGIN
Query Match 1.9%; Score 53; DB 9; Length 521;
Best Local Similarity 100.0%; Pred. No. 0.0041;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2723 AAACCTCTTAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2775
|||||.....
Db 426 AAACCTCTTAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 478

RESULT 12
BG632911
LOCUS
DEFINITION  BG632911 208 bp mRNA linear EST 23-APR-2001
melanogaster cDNA clone GH15835 5, mRNA sequence.
ACCESSION  BG632911
VERSION    BG632911.1 GI:13758395
KEYWORDS   EST.
SOURCE     fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
            1 (bases 1 to 208)
            Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
            BDGP/HMI Drosophila EST Project
            Unpublished (2001)
            Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: GH.158 row: C column: 11
            High quality sequence stop: 206.
            Location/Qualifiers
                1..208
                /organism="Drosophila melanogaster"
                /db_xref="taxon:7227"
                /clone="GH15835"
                /clone_lib="GH Drosophila melanogaster head pot2"
                /sex="male and female"
                /dev_stage="adult"
                /lab_host="DH5-alpha"
                /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                XhoI; Sized fractionated cDNAs were directly ligated into
                pOT2. Plasmid cDNA library."

```

```

BASE COUNT      107 a      18 c      21 g      62 t
ORIGIN

Query Match      1.9%; Score 52; DB 12; Length 208;
Best Local Similarity 100.0%; Pred. No. 0.017; 0; Indels 0; Gaps 0;
Matches 52; Conservative 0; Mismatches 0;

QY 2724 AACTCTTAAAAA..... 260 bp mRNA linear EST 03-DEC-2001
|||||.....
Db 146 AACTCTTAAAAA..... 260 bp mRNA linear EST 03-DEC-2001
|||||.....

RESULT 13
BM154294
LOCUS      fv83d01.y1 Zebrafish SJD adult male II Danio rerio cDNA clone
DEFINITION      5543832 5', mRNA sequence.
ACCESSION      BM154294
VERSION      BM154294.1 GI:17236260
KEYWORDS      EST.
SOURCE      zebrafish.
ORGANISM      Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
REFERENCE      1 (bases 1 to 260)
AUTHORS      Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
and Wilson,R.
TITLE      WashU Zebrafish EST Project 1998
JOURNAL      Unpublished (1998)
COMMENT      Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
cDNA Library construction by: Joe Barnes and Steve Johnson. DNA
Sequencing by: Washington University Genome Sequencing Center Clone
Distribution: Research Genetics web address:
http://www.researchgenetics.com/
Seq primer: T3 ET from Amersham
High quality sequence stop: 243.

FEATURES
Location/Qualifiers
1..260
/organism="Danio rerio"
/strain="SJD"
/db_xref="taxon:7955"
/clone_lib="5543832"
/clone="5543832"
/tissue_type="zebrafish SJD adult male II"
/sex="male"
/tissue_type="whole body"
/lab_host="DH10B"
/notes="Vector: pAMP1; Site_1: EcoRI; Site_2: NotI; First
strand cDNA synthesis was primed using oligo-dT on
magnetic beads with an additional primer
5'-ggcgcgctaatcactactacta-taggg-3'. Second strand
synthesis was a 3-cycle PCR using the primers
5'-ggcgctaatcactactacta-taggg-3' and
5'-aagcagtggtacacgcagtagcttt-tttttttttt-3'. The
cDNA was subsequently amplified in a 7-cycle PCR with the
following primers: 5'-ggcgctaatcactactacta-taggg-3' and
5'-aagcagtggtacacgcagtagcttt-tttttttttt-3'. The
cDNA was added in
a third PCR (5 cycles) and the primers
5'-caucacacgagggcgctaatcactactacta-taggg-3' and
5'-cuacacacgagggcgctaatcactactacta-taggg-3'. Ends were
treated with uracil DNA glycosylase and product with 3'
overhangs was annealed to complementary ends of pAMP1.
Insert can be excised using EcoRI and NotI. Library
constructed by Joe Barnes and Steve Johnson (Washington

```

```

BASE COUNT      139 a      29 c      15 g      77 t
ORIGIN

Query Match      1.9%; Score 52; DB 13; Length 260;
Best Local Similarity 100.0%; Pred. No. 0.014; 0; Indels 0; Gaps 0;
Matches 52; Conservative 0; Mismatches 0;

QY 2724 AACTCTTAAAAA..... 354 bp mRNA linear EST 07-FEB-1998
|||||.....
Db 175 AACTCTTAAAAA..... 354 bp mRNA linear EST 07-FEB-1998
|||||.....

RESULT 14
AA761919/c
LOCUS      nz42h05.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1290489 3',
DEFINITION      mRNA sequence.
ACCESSION      AA761919
VERSION      AA761919
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 354)
AUTHORS      NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1207 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 258.

FEATURES
Location/Qualifiers
1..354
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1290489"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAAGTGGAGCGGCGCTCATTTTTTTTTTTT-3'
]. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT      99 a      53 c      69 g      133 t
ORIGIN

Query Match      1.9%; Score 52; DB 9; Length 354;
Best Local Similarity 100.0%; Pred. No. 0.01; 0; Indels 0; Gaps 0;
Matches 52; Conservative 0; Mismatches 0;

QY 2724 AACTCTTAAAAA..... 133 t
|||||.....

```



Db 57 AACCTTAA 6

RESULT 15  
BE198722 358 bp mRNA linear EST 26-JUN-2000  
LOCUS u981g09.y1 Soares\_mammary\_gland\_NMLMG Mus musculus cDNA clone  
DEFINITION IMAGE:1548832 5', mRNA sequence.  
ACCESSION BE198722  
VERSION BE198722.1 GI:8710891  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 358)  
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
TITLE Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
This clone is available royalty-free through LNL ; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
MGI:954180  
Seq primer: -40RP from Gibco  
High quality sequence stop: 347.

FEATURES  
source  
1..358  
Location/Qualifiers  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:1548832"  
/clone\_lib="Soares\_mammary\_gland\_NMLMG"  
/sex="female (lactating)"  
/tissue\_type="mammary gland"  
/lab\_host="DH10B"  
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified  
polylinker; 1st strand cDNA was prepared from mammary  
gland tissue from a lactating female, and was then primed  
with a Not I - oligo(dT) primer. Double-stranded cDNA was  
ligated to Eco RI adaptors (Pharmacia), digested with Not  
I and cloned into the Not I and Eco RI sites of the  
modified pT7T3 vector. Library is normalized. Library  
was constructed by Bento Soares and M. Fatima Bonaldo."  
BASE COUNT 136 a 75 c 82 g 65 t  
ORIGIN

Query Match 1.9%; Score 52; DB 10; Length 358;  
Best Local Similarity 100.0%; Pred. No. 0.0099;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2724 AACCTTAA 2775  
Db 295 AACCTTAA 346

Search completed: April 12, 2003, 08:58:50  
Job time : 3061.33 secs

GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 12, 2003, 03:59:25 ; Search time 77.0226 Seconds  
(without alignments)  
11049.076 Million cell updates/sec

Title: US-09-914-152-2  
Perfect score: 2775  
Sequence: 1 gtaattctctttctgtc.....aaaaaaaaaaaaaaaaaaaaa 2775

Scoring table: OLIGO-NUC  
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A-COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B-COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A-COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B-COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PctUS-COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	50	1.8	1069	4 US-09-372-422A-7
2	50	1.8	1474	4 US-08-821-994-64
3	50	1.8	2301	4 US-09-232-191-8
4	50	1.8	2301	4 US-09-232-200-8
5	50	1.8	2301	4 US-09-232-197-8
6	50	1.8	2301	4 US-09-232-201-8
7	50	1.8	2710	4 US-09-232-200-44
8	50	1.8	2710	4 US-09-232-200-70
9	50	1.8	2710	4 US-09-232-197-44
10	50	1.8	2710	4 US-09-232-197-70
11	50	1.8	2710	4 US-09-232-201-44
12	50	1.8	2710	4 US-09-232-201-70
13	49	1.8	1159	4 US-09-410-464-14
14	49	1.8	1490	2 US-08-553-367A-5
15	49	1.8	1490	4 US-09-295-306-5
16	49	1.8	1490	4 US-09-734-719-5
17	48	1.7	1737	1 US-08-202-056-4
18	48	1.7	1737	1 US-08-076-093A-3
19	48	1.7	1737	1 US-08-701-265-3
20	48	1.7	1737	2 US-08-284-586-3
21	48	1.7	1737	2 US-08-805-478-3
22	48	1.7	1737	2 US-08-802-627A-3
23	48	1.7	1737	2 US-08-801-238-3
24	48	1.7	1737	2 US-08-801-228-3
25	48	1.7	1737	3 US-09-104-296-3
26	48	1.7	1737	5 US-09-104-296-3
27	48	1.7	1998	4 US-09-232-200-68

28	48	1.7	1998	4	US-09-232-197-68	Sequence 68, Appl
29	48	1.7	1998	4	US-09-232-201-68	Sequence 68, Appl
30	48	1.7	2087	4	US-09-232-191-6	Sequence 6, Appl
31	48	1.7	2087	4	US-09-232-200-6	Sequence 6, Appl
32	48	1.7	2087	4	US-09-232-197-6	Sequence 6, Appl
33	48	1.7	2087	4	US-09-232-201-6	Sequence 6, Appl
34	48	1.7	2378	4	US-08-802-805D-20	Sequence 20, Appl
35	48	1.7	2719	3	US-08-706-216-1	Sequence 1, Appl
C 37	47	1.7	80	4	US-09-284-627-15	Sequence 15, Appl
C 38	47	1.7	105	4	US-09-284-627-23	Sequence 23, Appl
39	47	1.7	120	1	US-08-153-051B-28	Sequence 28, Appl
40	47	1.7	120	1	US-08-060-952C-44	Sequence 44, Appl
41	47	1.7	120	2	US-08-151-477A-28	Sequence 28, Appl
42	47	1.7	120	3	US-08-819-867-58	Sequence 58, Appl
43	47	1.7	120	4	US-08-464-011B-44	Sequence 44, Appl
44	47	1.7	144	1	US-08-702-344-26	Sequence 26, Appl
45	47	1.7	222	4	US-08-481-190-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1  
US-09-372-422A-7  
; Sequence 7, Application US/09372422A  
; Patent No. 6313375  
; GENERAL INFORMATION:  
; APPLICANT: Rudolf Jung  
; TITLE OF INVENTION: Maize Aquaporins and Uses Thereof  
; FILE REFERENCE: 0919  
; CURRENT APPLICATION NUMBER: US/09/372.422A  
; PRIOR FILING DATE: 1999-08-11  
; PRIOR APPLICATION NUMBER: US 60/098,692  
; PRIOR FILING DATE: 1998-08-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7  
; LENGTH: 1069  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (100)...(946)  
US-09-372-422A-7  
  
Query Match 1.8%; Score 50; DB 4; Length 1069;  
Best Local Similarity 100.0%; Pred. No. 2.9e-11;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2726 CTTTAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2775  
|||||.....  
DB 973 CTTTAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1022  
|||||.....

RESULT 2  
US-08-821-994-64  
; Sequence 64, Application US/08821994A  
; Patent No. 6228643  
; GENERAL INFORMATION:  
; APPLICANT: Greenland, Andrew J  
; APPLICANT: Thomas, Didier Rp  
; APPLICANT: Jepson, Ian  
; TITLE OF INVENTION: Promoters  
; FILE REFERENCE: PPD 50108  
; CURRENT APPLICATION NUMBER: US/08/821.994A  
; CURRENT FILING DATE: 1997-03-22  
; EARLIER APPLICATION NUMBER: PCT/GB97/00729  
; EARLIER FILING DATE: 1997-03-18  
; EARLIER APPLICATION NUMBER: GB 9606062.9  
; EARLIER FILING DATE: 1996-03-22  
; NUMBER OF SEQ ID NOS: 89

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Brassica napus
US-08-821-994-64

```

```

Query Match      1.8%; Score 50; DB 4; Length 1474;
Best Local Similarity 100.0%; Pred. No. 2.8e-11;
Matches 50; Conservative 0; Mismatches 0; Indels
QY 2726 CTCCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2775
Db 1371 CTCCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1420

```

### RESULT 3

```

RESULT 3
US-09-232-191-8
; Sequence 8, Application US/09232191
; Patent No. 6284487
; GENERAL INFORMATION:
; APPLICANT: Stahl, Andreas
; APPLICANT: Hirsch, David J.
; APPLICANT: Lodish, Harvey F.
; TITLE OF INVENTION: Fatty Acid Transport Proteins
; FILE REFERENCE: WH197-2193ME
; CURRENT APPLICATION NUMBER: US/09/232.191
; CURRENT FILING DATE: 1999-01-14
; EARLIER APPLICATION NUMBER: 60/071,374
; EARLIER FILING DATE: 1998-01-15
; EARLIER APPLICATION NUMBER: 60/093,491
; EARLIER FILING DATE: 1998-07-20
; EARLIER APPLICATION NUMBER: 60/110,941
; EARLIER FILING DATE: 1998-12-04
; NUMBER OF SEQ IDS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-232-191-8

```

```

Query Match      1.8%; Score 50; DB 4; Length 2301;
Best Local Similarity 100.0%; Pred. No. 2.6e-11;
Matches 50; Conservative 0; Mismatches 0; Indels 0
QY 2726 CTCCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2775
Db 2240 CTCCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2289

```

## RESULT 4

```

US/09-232-200-8
: Sequence 8, Application US/09232200A
: Patent No. 6388213
: GENERAL INFORMATION:
: APPLICANT: Stahl, Andreas
: APPLICANT: Hirsch, David J.
: APPLICANT: Lodish, Harvey F.
: APPLICANT: Gimeno, Ruth E.
: APPLICANT: Tartaglia, Louis A.
: TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
: FILE REFERENCE: WHI97-21p3MB
: CURRENT APPLICATION NUMBER: US/09/232,200A
: CURRENT FILING DATE: 1999-01-14
: EARLIER APPLICATION NUMBER: 60/071,374
: EARLIER FILING DATE: 1998-01-15
: EARLIER APPLICATION NUMBER: 60/093,491
: EARLIER FILING DATE: 1998-07-20
: EARLIER APPLICATION NUMBER: 60/110,941
: EARLIER FILING DATE: 1998-12-04
: NUMBER OF SEQ IDS NOS: 105
: SOFTWARE: FastSeq for Windows Version 3.0

```

```

; SEQ ID NO 8
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-232-200-8

```

```

Query Match      1.9%; Score 50; DB 4; Length 2301;
Best Local Similarity 100.0%; Pred. No. 2.ee-11;
Matches 50; Conservative 0; Mismatches 0; Indels

QY 2726 CTCCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2775
Db 2240 CTCCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2289

```

## RESULT 5

```

US-09-232-197-8
; Sequence 8, Application US/09232197A
; Patent No. 6300096
; GENERAL INFORMATION:
; APPLICANT: Stahl, Andreas
; APPLICANT: Hirsch, David J.
; APPLICANT: Lodish, Harvey F.
; APPLICANT: Gimeno, Ruth E.
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
; FILE REFERENCE: WH197-21D3MA
; CURRENT APPLICATION NUMBER: US/09/232,197A
; CURRENT FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/071,374
; EARLIER FILING DATE: 1998-01-15
; EARLIER APPLICATION NUMBER: 60/093,491
; EARLIER FILING DATE: 1998-07-20
; EARLIER APPLICATION NUMBER: 60/110,941
; EARLIER FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-232-197-8

```

Query Match 1.8%; Score 50; DB 4; Length 2301;  
Best Local Similarity 100.0%; Pred. No. 2.e-11;  
Matches 50; Conservative 0; Mismatches 0; Indels

QY 2726 TCCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2775  
Db 2240 TCCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2289

## RESULT 6

```

US-09-232-201-8
; Sequence 8, Application US/09232201A
; Patent No. 6348321
; GENERAL INFORMATION:
; APPLICANT: Stahl, Andreas
; APPLICANT: Hirsch, David J.
; APPLICANT: Lodish, Harvey F.
; APPLICANT: Gimeno, Ruth E.
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
; FILE REFERENCE: WHI97-21p3MC
; CURRENT APPLICATION NUMBER: US/09/232,201A
; CURRENT FILING DATE: 1999-01-14
; EARLIER APPLICATION NUMBER: 60/071,374
; EARLIER FILING DATE: 1998-01-15
; EARLIER APPLICATION NUMBER: 60/093,491
; EARLIER FILING DATE: 1998-07-20
; EARLIER APPLICATION NUMBER: 60/110,941
; EARLIER FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 105

```

```
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-232-201-8

Query Match      1.8%; Score 50; DB 4; Length 2301;
Best Local Similarity 100.0%; Pred. No. 2.6e-11;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2726 CTCCTTAAAAA 18
Db 2240 CTCCTTAAAAA 18

RESULT 7
US-09-232-200-44
; Sequence 44, Application US/09232200A
; Patent No. 6288213
; GENERAL INFORMATION:
; APPLICANT: Stahl, Andreas
; APPLICANT: Hirsch, David J.
; APPLICANT: Lodish, Harvey F.
; APPLICANT: Gimeno, Ruth E.
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
; FILE REFERENCE: WHI97-21p3MB
; CURRENT APPLICATION NUMBER: US/09/232,200A
; CURRENT FILING DATE: 1999-01-14
; EARLIER APPLICATION NUMBER: 60/071,374
; EARLIER FILING DATE: 1998-01-15
; EARLIER APPLICATION NUMBER: 60/093,491
; EARLIER FILING DATE: 1998-07-20
; EARLIER APPLICATION NUMBER: 60/110,941
; EARLIER FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 44
; LENGTH: 2710
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-232-200-44

Query Match      1.8%; Score 50; DB 4; Length 2710;
Best Local Similarity 100.0%; Pred. No. 2.6e-11;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2726 CTCCTTAAAAA 18
Db 2649 CTCCTTAAAAA 18

RESULT 8
US-09-232-200-70
; Sequence 70, Application US/09232200A
; Patent No. 6288213
; GENERAL INFORMATION:
; APPLICANT: Stahl, Andreas
; APPLICANT: Hirsch, David J.
; APPLICANT: Lodish, Harvey F.
; APPLICANT: Gimeno, Ruth E.
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
; FILE REFERENCE: WHI97-21p3MB
; CURRENT APPLICATION NUMBER: US/09/232,200A
; CURRENT FILING DATE: 1999-01-14
; EARLIER APPLICATION NUMBER: 60/071,374
; EARLIER FILING DATE: 1998-01-15
; EARLIER APPLICATION NUMBER: 60/093,491
; EARLIER FILING DATE: 1998-07-20
; EARLIER APPLICATION NUMBER: 60/110,941
; EARLIER FILING DATE: 1998-12-04
```

```
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 70
; LENGTH: 2710
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-232-200-70

Query Match      1.8%; Score 50; DB 4; Length 2710;
Best Local Similarity 100.0%; Pred. No. 2.6e-11;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2726 CTCCTTAAAAA 18
Db 2649 CTCCTTAAAAA 18

RESULT 9
US-09-232-197-44
; Sequence 44, Application US/09232197A
; Patent No. 6300096
; GENERAL INFORMATION:
; APPLICANT: Stahl, Andreas
; APPLICANT: Hirsch, David J.
; APPLICANT: Lodish, Harvey F.
; APPLICANT: Gimeno, Ruth E.
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
; FILE REFERENCE: WHI97-21p3MA
; CURRENT APPLICATION NUMBER: US/09/232,197A
; CURRENT FILING DATE: 1999-01-14
; EARLIER APPLICATION NUMBER: 60/071,374
; EARLIER FILING DATE: 1998-01-15
; EARLIER APPLICATION NUMBER: 60/093,491
; EARLIER FILING DATE: 1998-07-20
; EARLIER APPLICATION NUMBER: 60/110,941
; EARLIER FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 44
; LENGTH: 2710
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-232-197-44

Query Match      1.8%; Score 50; DB 4; Length 2710;
Best Local Similarity 100.0%; Pred. No. 2.6e-11;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2726 CTCCTTAAAAA 18
Db 2649 CTCCTTAAAAA 18

RESULT 10
US-09-232-197-70
; Sequence 70, Application US/09232197A
; Patent No. 6300096
; GENERAL INFORMATION:
; APPLICANT: Stahl, Andreas
; APPLICANT: Hirsch, David J.
; APPLICANT: Lodish, Harvey F.
; APPLICANT: Gimeno, Ruth E.
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
; FILE REFERENCE: WHI97-21p3MA
; CURRENT APPLICATION NUMBER: US/09/232,197A
; CURRENT FILING DATE: 1999-01-14
; EARLIER APPLICATION NUMBER: 60/071,374
; EARLIER FILING DATE: 1998-01-15
; EARLIER APPLICATION NUMBER: 60/093,491
; EARLIER FILING DATE: 1998-07-20
; EARLIER APPLICATION NUMBER: 60/110,941
; EARLIER FILING DATE: 1998-12-04
```

; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 70  
; LENGTH: 2710  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-232-197-70

Query Match 1.8%; Score 50; DB 4; Length 2710;  
Best Local Similarity 100.0%; Pred. No. 2.6e-11;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2726 CTCCTTAAAAA... 2775  
Db 2649 CTCCTTAAAAA... 2698

RESULT 11  
US-09-232-201-44  
; Sequence 44, Application US/09232201A  
; Patent No. 6348321  
; GENERAL INFORMATION:  
; APPLICANT: Stahl, Andreas  
; APPLICANT: Hirsch, David J.  
; APPLICANT: Lodish, Harvey F.  
; APPLICANT: Gimeno, Ruth E.  
; APPLICANT: Tartaglia, Louis A.  
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
; FILE REFERENCE: WHI97-21p3MC  
; CURRENT APPLICATION NUMBER: US/09/232,201A  
; CURRENT FILING DATE: 1999-01-14  
; EARLIER APPLICATION NUMBER: 60/071,374  
; EARLIER FILING DATE: 1998-01-15  
; EARLIER APPLICATION NUMBER: 60/093,491  
; EARLIER FILING DATE: 1998-07-20  
; EARLIER APPLICATION NUMBER: 60/110,941  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 44  
; LENGTH: 2710  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-232-201-44

Query Match 1.8%; Score 50; DB 4; Length 2710;  
Best Local Similarity 100.0%; Pred. No. 2.6e-11;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2726 CTCCTTAAAAA... 2775  
Db 2649 CTCCTTAAAAA... 2698

RESULT 12  
US-09-232-201-70  
; Sequence 70, Application US/09232201A  
; Patent No. 6348321  
; GENERAL INFORMATION:  
; APPLICANT: Stahl, Andreas  
; APPLICANT: Hirsch, David J.  
; APPLICANT: Lodish, Harvey F.  
; APPLICANT: Gimeno, Ruth E.  
; APPLICANT: Tartaglia, Louis A.  
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
; FILE REFERENCE: WHI97-21p3MC  
; CURRENT APPLICATION NUMBER: US/09/232,201A  
; CURRENT FILING DATE: 1999-01-14  
; EARLIER APPLICATION NUMBER: 60/071,374  
; EARLIER FILING DATE: 1998-01-15  
; EARLIER APPLICATION NUMBER: 60/093,491  
; EARLIER FILING DATE: 1998-07-20

; EARLIER APPLICATION NUMBER: 60/110,941  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 70  
; LENGTH: 2710  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-232-201-70

Query Match 1.8%; Score 50; DB 4; Length 2710;  
Best Local Similarity 100.0%; Pred. No. 2.6e-11;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2726 CTCCTTAAAAA... 2775  
Db 2649 CTCCTTAAAAA... 2698

RESULT 13  
US-09-410-464-14  
; Sequence 14, Application US/09410464  
; Patent No. 6395892  
; GENERAL INFORMATION:  
; APPLICANT: Strauss et al.  
; TITLE OF INVENTION: Floral homeotic genes for manipulation of flowering in  
; TITLE OF INVENTION: poplar and other plant species.  
; FILE REFERENCE: 53375  
; CURRENT APPLICATION NUMBER: US/09/410,464  
; CURRENT FILING DATE: 1999-10-01  
; EARLIER APPLICATION NUMBER: 09/287,700  
; EARLIER FILING DATE: 1999-04-06  
; EARLIER APPLICATION NUMBER: 60/080,851  
; EARLIER FILING DATE: 1998-04-06  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 1159  
; TYPE: DNA  
; ORGANISM: Populus balsamifera subsp. trichocarpa  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (99)...(815)  
US-09-410-464-14

Query Match 1.8%; Score 49; DB 4; Length 1159;  
Best Local Similarity 100.0%; Pred. No. 7.2e-11;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2727 TCTTAAAAA... 2775  
Db 1108 TCTTAAAAA... 1156

RESULT 14  
US-08-553-367A-5  
; Sequence 5, Application US/08553367A  
; Patent No. 5939539  
; GENERAL INFORMATION:  
; APPLICANT: Theodor LANGE et al.  
; TITLE OF INVENTION: GA 20-OXIDASE GENE SEQUENCES (As Amended)  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.  
; STREET: 2033 K Street, N.W., Suite 800  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS

SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/553,367A  
FILING DATE: No. 5939539ember 27, 1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee Cheng  
REGISTRATION NUMBER: 40,949  
REFERENCE/DOCKET NUMBER: 49/FD4.5M2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-721-8200  
TELEFAX: 202-721-8250  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1490 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Arabidopsis thaliana  
IMMEDIATE SOURCE:  
CLONE: PAT2353  
US-08-553-367A-5

Query Match 1.8%; Score 49; DB 2; Length 1490;  
Best Local Similarity 100.0%; Pred. No. 7e-11;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2727 TCTTAAAAA  
|||||  
DB 1420 TCTTAAAAA

RESULT 15  
US-09-295-306-5  
Sequence 5, Application US/09295306  
Patent No. 6138021  
GENERAL INFORMATION:  
APPLICANT: Theodor LANGE et al.  
TITLE OF INVENTION: GA 20-OXIDASE GENE SEQUENCES (As Amended)  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Wenderoth, Lind & Ponack, L.L.P.  
STREET: 2033 K Street, N.W., Suite 800  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/295,306  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/553,367  
FILING DATE: No. 6198021ember 27, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee Cheng  
REGISTRATION NUMBER: 40,949  
REFERENCE/DOCKET NUMBER: 49/DIV-FD4.5M2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-721-8200

TELEFAX: 202-721-8250  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1490 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Arabidopsis thaliana  
IMMEDIATE SOURCE:  
CLONE: PAT2353  
US-09-295-306-5  
Query Match 1.8%; Score 49; DB 4; Length 1490;  
Best Local Similarity 100.0%; Pred. No. 7e-11;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2727 TCTTAAAAA  
|||||  
DB 1420 TCTTAAAAA  
Search completed: April 12, 2003, 09:02:18  
Job time : 106.023 secs



```

; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 1443
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-904-615-21

Query Match 1.8%; Score 50; DB 10; Length 1443;
Best Local Similarity 100.0%; Pred. No. 2e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2726 CTCCTTAAAAA 100.0%
Db 1387 CTCCTTAAAAA 100.0%

RESULT 3
US-09-764-864-239
; Sequence 239, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 239
; LENGTH: 1801
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-239

Query Match 1.8%; Score 50; DB 10; Length 1801;
Best Local Similarity 100.0%; Pred. No. 2e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2726 CTCCTTAAAAA 100.0%
Db 1725 CTCCTTAAAAA 100.0%

RESULT 4
US-10-125-459-2
; Sequence 2, Application US/10125459
; Publication No. US20020192800A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Serine proteases
; FILE REFERENCE: PT005P1
; CURRENT APPLICATION NUMBER: US/10/125,459
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US/09/597,839
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 60/133,239
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 60/135,163
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 60/147,005
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/152,935
```

```

; PRIOR FILING DATE: 1999-09-09
; PRIOR APPLICATION NUMBER: 60/162,979
; PRIOR FILING DATE: 1999-11-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1939
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (308)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1915)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1938)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-125-459-2

Query Match 1.8%; Score 50; DB 9; Length 1939;
Best Local Similarity 100.0%; Pred. No. 2e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2726 CTCCTTAAAAA 100.0%
Db 1853 CTCCTTAAAAA 100.0%

RESULT 5
US-10-067-761-4
; Sequence 4, Application US/10067761
; Publication No. US20020197701A1
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: Serine protease Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PT005P4
; CURRENT APPLICATION NUMBER: US/10/067,761
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 09/804,156
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/189,025
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1939
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (308)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1915)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1938)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-067-761-4

Query Match 1.8%; Score 50; DB 9; Length 1939;
Best Local Similarity 100.0%; Pred. No. 2e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2726 CTCCTTAAAAA 100.0%
Db 1853 CTCCTTAAAAA 100.0%
```





```

; NAME/KEY: misc-feature
; LOCATION: (437)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-091-504-332

Query Match      1.8%; Score 49; DB 9; Length 439;
Best Local Similarity 100.0%; Pred. No. 6.7e-15;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2727 TCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2775
      |||||||
Db 378 TCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 426
      |||||||

RESULT 12
US-09-764-869-332
; Sequence 332, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 332
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (437)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-869-332

Query Match      1.8%; Score 49; DB 10; Length 439;
Best Local Similarity 100.0%; Pred. No. 6.7e-15;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2727 TCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2775
      |||||||
Db 378 TCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 426
      |||||||

RESULT 13
US-09-770-444-872/c
; Sequence 872, Application US/09770444
; Patent No. US2002003280A1
; GENERAL INFORMATION:
; APPLICANT: Goriach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krickler, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hofman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2002/0001016

```

; CURRENT APPLICATION NUMBER: US/09/770,444  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,502  
; PRIOR FILING DATE: 2000-01-27  
; NUMBER OF SEQ ID NOS: 999  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 872  
; LENGTH: 442  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-770-444-872

Query Match 1.8%; Score 49; DB 10; Length 442;  
Best Local Similarity 100.0%; Pred. No. 6.7e-15;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2727 TCTTAA 2775  
Db 67 TCTTAA 19

RESULT 14  
US-10-104-580-14  
; Sequence 14, Application US/10104580  
; Publication No. US2003003628A1  
; GENERAL INFORMATION:  
; APPLICANT: Strauss et al.  
; TITLE OF INVENTION: Floral homeotic genes for manipulation of flowering in  
; TITLE OF INVENTION: poplar and other plant species.  
; FILE REFERENCE: 62486  
; CURRENT APPLICATION NUMBER: US/10/104,580  
; CURRENT FILING DATE: 2002-03-21  
; PRIOR APPLICATION NUMBER: 09/410,464  
; PRIOR FILING DATE: 1999-10-01  
; PRIOR APPLICATION NUMBER: 09/287,700  
; PRIOR FILING DATE: 1999-04-06  
; PRIOR APPLICATION NUMBER: 60/080,851  
; PRIOR FILING DATE: 1998-04-06  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 1159  
; TYPE: DNA  
; ORGANISM: Populus balsamifera subsp. trichocarpa  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (99)..(815)  
US-10-104-580-14

Query Match 1.8%; Score 49; DB 9; Length 1159;  
Best Local Similarity 100.0%; Pred. No. 6.6e-15;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2727 TCTTAA 2775  
Db 1108 TCTTAA 1156

RESULT 15  
US-09-925-301-545  
; Sequence 545, Application US/09925301  
; Patent No. US20020052308A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA106  
; CURRENT APPLICATION NUMBER: US/09/925,301  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05882  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1694

; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 545  
; LENGTH: 1176  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-301-545

Query Match 1.8%; Score 49; DB 10; Length 1176;  
Best Local Similarity 100.0%; Pred. No. 6.6e-15;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2727 TCTTAA 2775  
Db 1127 TCTTAA 1175

Search completed: April 12, 2003, 12:47:52  
Job time : 213.677 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 14:22:18 ; Search time 1364.65 Seconds  
(without alignments)  
15994.713 Million cell updates/sec

Title: US-09-914-152-3\_COPY\_1\_750

Perfect score: 750

Sequence: 1 cgcctctgcaaggtagacc.....ctagaatagaattatcgt 750

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vl.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vl.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_nam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	750	100.0	10562	6	E38420	E38420 Novel polyp
2	750	100.0	170121	9	AF064860	AF064860 Homo sapi
3	750	100.0	340000	9	HS21C080	AL163280 Homo sapi
4	483.2	64.4	185982	2	AC073231	AC073231 Homo sapi
c	135	18.0	185982	2	AC073231	AC073231 Homo sapi
5	61.6	8.2	22295	9	AC007088	AC007088 Homo sapi
c	57.8	7.7	162063	9	AC015553	AC015553 Homo sapi
c	57.8	7.7	176932	9	AL157884	AL157884 Human DNA
8	56.6	7.5	118819	9	AP000647	AP000647 Homo sapi
c	56.6	7.5	173359	2	AP000922	AP000922 Homo sapi
11	55.4	7.4	68444	2	AC114686	AC114686 Homo sapi
c	55.4	7.4	188723	9	CNS01DU9	AL133224 Human Chr
12	55.4	7.4	204634	2	AC007346	AC007346 Homo sapi
c	55.2	7.4	190565	9	AC007374	AC007374 Homo sapi
15	55	7.3	207558	2	AC114778	AC114778 Homo sapi
c	54.4	7.3	157402	2	AC034161	AC034161 Homo sapi
17	54	7.2	94359	9	AP000974	AP000974 Homo sapi
c	54	7.2	195760	2	AC016486	AC016486 Homo sapi
19	53.8	7.2	119944	9	AC111162	AC111162 Homo sapi
20	53.8	7.2	159723	2	AC021862	AC021862 Homo sapi
c	53.8	7.2	165287	9	AC090599	AC090599 Homo sapi
22	53.8	7.2	166240	2	AC110599	AC110599 Homo sapi
c	53.6	7.1	73334	6	AX347027	AX347027 Sequence
c	53.6	7.1	73334	6	AX358494	AX358494 Sequence
c	53.6	7.1	122427	9	AC025459	AC025459 Homo sapi
26	53.4	7.1	4601	3	DMU11584	U11584 Drosophila
27	53.4	7.1	6593	9	AL513367	AL513367 Human DNA
28	53.4	7.1	19517	3	DMG37541	U37541 Drosophila
29	53.4	7.1	185893	9	AL133383	AL133383 Human DNA
c	53.4	7.1	204158	9	AL133383	AL133383 Human DNA
31	53	7.1	23760	9	AL450351	AL450351 Homo sapi
c	52.8	7.0	163903	9	AC007387	AC007387 Homo sapi
c	52.8	7.0	214227	2	AL451003	AL451003 Homo sapi
c	52.8	7.0	234523	9	AC006544	AC006544 Homo sapi
c	52.6	7.0	144201	9	AC112213	AC112213 Homo sapi
c	52.6	7.0	174293	9	AC113169	AC113169 Homo sapi
c	52.4	7.0	99995	9	AC091157	AC091157 Homo sapi
c	52.4	7.0	178905	9	AC079316	AC079316 Homo sapi
c	52.4	7.0	192239	9	AC009796	AC009796 Homo sapi
40	52.4	7.0	204140	9	AC010976	AC010976 Homo sapi
41	52.4	7.0	205342	9	AC068282	AC068282 Homo sapi
c	52.4	7.0	219973	2	AC092906	AC092906 Homo sapi
c	52.2	7.0	39163	9	HSSYBL16	AJ225782 Homo sapi
c	52.2	7.0	113639	9	HSJ744A17	AL096862 Human DNA
45	52.2	7.0	158661	9	HSA271736	AJ271736 Homo sapi

ALIGNMENTS

RESULT 1	E38420	Novel polypeptide.	10562 bp	DNA	linear	PAT 31-JAN-2002
E38420	LOCUS					
DEFINITION	E38420	Novel polypeptide.				
ACCESSION	E38420.1	GI:18626994				
VERSION	JP 2000245464-A/2.					
KEYWORDS	Homo sapiens.					
SOURCE	Homo sapiens					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 10562)					
AUTHORS	Narimatsu, H., Isshiki, S., Togayauchi, A. and Sasaki, K.					
TITLE	Novel polypeptide					
JOURNAL	Patent: JP 2000245464-A 2 12-SEP-2000;					





```

/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="BAC:291B3, 3' partial"
/clone_lib="BAC library"
/note="Accession No. AF064857"
1088..1566
/note="L1M4"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
2398..2495
/note="(TCTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
complement(2496..2878)
/note="THEIC"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
2879..2931
/note="(TCTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
2932..3004
/note="(TCCA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
3005..3074
/note="(TCCA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
3473..19157
/gene="SH3BGR"
join(<3473..3565,12649..12678,15331..15460,18739..19157)
/gene="SH3BGR"
join(3473..3565,12649..12678,15331..15426)
/partial
/gene="SH3BGR"
/note="Accession No. X93498"
/codon_start=1
/product="21-Glutamic Acid Rich protein 21-GARP"
/db_xref="GI:7171370"
/translation="GSEKAEKGTEAQKESGVNLPQAEKNEEGETATEETE
IAMGAGEAEETABEGEPGEDS"
3473..3565
/gene="SH3BGR"
/number=4
5542..5829
/note="AluSx"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(6887..7067)
/note="L1MD3"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
complement(8482..8762)
/note="AluJo"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(8865..8964)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
9131..9424
/note="AluSq"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(9723..9882)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
complement(9989..10077)
/note="L2"

/rpt_family="LINE/L2"
/rpt_type=DISPERSED
complement(10078..10429)
/note="THE1B"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
complement(10430..10651)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
10980..11000
/note="(TTTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
complement(11002..11282)
/note="AluSp"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(11315..11556)
/note="AluJo"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
11702..11753
/note="MIR"
/rpt_family="SINE/MIR"
/rpt_type=DISPERSED
12649..12678
/gene="SH3BGR"
/number=5
complement(13115..13425)
/note="AluSx"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(13598..13892)
/note="AluY"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
14013..14131
/note="(CA)n"
/rpt_family="Simple_repeat"

Query Match 100.0%; Score 750; DB 9; Length 340000;
Best Local Similarity 100.0%; Pred. No. 7.4e-169;
Matches 750; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCCTCTGGCAAGGTAGACCTTGAAGCAAACTGAGTTGAGTTGTTAGACGGAAT 60
Db 155223 CGCCTCTGGCAAGGTAGACCTTGAAGCAAACTGAGTTGAGTTGTTAGACGGAAT 155282

Qy 61 AATTACTGTGGCATGCAGCACTTCCCAACCGTTCTGTGAGGAGGAGGAGTGTATTGCC 120
Db 155283 AATTACTGTGGCATGCAGCACTTCCCAACCGTTCTGTGAGGAGGAGGAGTGTATTGCC 155342

Qy 121 AGTTTGGCACAAAGGGCACAGGTGTAGAACACGTAAAGTCCCTGGCGCTGTACACCACC 180
Db 155343 AGTTTGGCACAAAGGGCACAGGTGTAGAACACGTAAAGTCCCTGGCGCTGTACACCACC 155402

Qy 181 ACTGTGTTTGTAGCTGAGATGTGAACCAAGGCGCTTCTGATTCCTCAAAATTCCTCATTCCTTTC 240
Db 155403 ACTGTGTTTGTAGCTGAGATGTGAACCAAGGCGCTTCTGATTCCTCAAAATTCCTCATTCCTTTC 155462

Qy 241 ATCTAGCAGGCTGCTCGGGTTAGCAGAAGGGACCTCTGTATCTGCTCTGCAGCTTCT 300
Db 155463 ATCTAGCAGGCTGCTCGGGTTAGCAGAAGGGACCTCTGTATCTGCTCTGCAGCTTCT 155522

Qy 301 TCAGCTGATTTATATGGAAACAGAGTAGATATTGATTTGGCAATTAGTGAATATTAT 360
Db 155523 TCAGCTGATTTATATGGAAACAGAGTAGATATTGATTTGGCAATTAGTGAATATTAT 155582

Qy 361 GAGATCATATAGCAAACTTCACAGTTTCATCAAGGATCCTGCCTTCAATATCTGGCCA 420
Db 155583 GAGATCATATAGCAAACTTCACAGTTTCATCAAGGATCCTGCCTTCAATATCTGGCCA 155642

```

```
QY 421 ACTGATGTGTAAGCAGCTGCAAGAACTTCAGAGCTCACAAAAAAGCAAACTCCAGAC 480
|||||
Db 155643 ACTGATGTGTAAGCAGCTGCAAGAACTTCAGAGCTCACAAAAAAGCAAACTCCAGAC 155702
|||||
QY 481 TTTATTTCCCTGGAACTCTGTTTGTGAGACACTGCCCAATGAATGCTCTCCAGAAATAG 540
|||||
Db 155703 TTTATTTCCCTGGAACTCTGTTTGTGAGACACTGCCCAATGAATGCTCTCCAGAAATAG 155762
|||||
QY 541 TCGGATTTGGTCAATAAATTTGGCAATTTACAGAACATGCTGCTTTTTCAGAGAT 600
|||||
Db 155763 TCGGATTTGGTCAATAAATTTGGCAATTTACAGAACATGCTGCTTTTTCAGAGAT 155822
|||||
QY 501 TTTATTTTAACTTAATTAATAATTAACATGATGATGATGATGATGATGATGATGATGAT 660
|||||
Db 155823 TTTATTTTAACTTAATTAATAATTAACATGATGATGATGATGATGATGATGATGAT 155882
|||||
QY 661 GTGCAATCTTCCTACATGCTCAATGAGTTTACATATATTTCCACCAATGATCA 720
|||||
Db 155883 GTGCAATCTTCCTACATGCTCAATGAGTTTACATATATTTCCACCAATGATCA 155942
|||||
QY 721 CCACCGAGATCTAGATAGATGATTCGT 750
|||||
Db 155943 CCACCGAGATCTAGATAGATGATTCGT 155972
|||||

RESULT 4
AC073231
LOCUS
DEFINITION Homo sapiens chromosome 17 clone RP11-729A4, WORKING DRAFT
ACCESSION AC073231
VERSION AC073231.1 GI:8440043
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 185982)
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 185982)
Waterston,R.H.
Direct Submission
Submitted (10-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

----- Genome Center -----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1724: contig of 1724 bp in length
* 1725 1824: gap of unknown length
* 1825 2971: contig of 1147 bp in length
* 2972 3071: gap of unknown length
* 3072 4660: contig of 1589 bp in length
* 4661 4760: gap of unknown length
* 4761 5930: contig of 1170 bp in length
* 5931 6030: gap of unknown length
* 6031 8127: contig of 2097 bp in length
* 8128 8228: gap of unknown length
* 8228 10160: contig of 1933 bp in length
* 10161 10260: gap of unknown length
* 10261 13236: contig of 2976 bp in length
* 13237 13336: gap of unknown length
* 13337 16603: contig of 3267 bp in length
* 16604 16703: gap of unknown length
*
* 20684: contig of 3981 bp in length
* 20784: gap of unknown length
* 24835: contig of 4051 bp in length
* 24936: gap of unknown length
* 29727: contig of 4792 bp in length
* 29828: gap of unknown length
* 34947: contig of 5120 bp in length
* 35048: gap of unknown length
* 41164: contig of 6117 bp in length
* 41264: gap of unknown length
* 47050: contig of 5786 bp in length
* 47051 47150: gap of unknown length
* 47151 53084: contig of 5934 bp in length
* 53184: gap of unknown length
* 53185 60624: contig of 7440 bp in length
* 60625 60724: gap of unknown length
* 60725 69057: contig of 8333 bp in length
* 69058 69157: gap of unknown length
* 69158 78529: contig of 9372 bp in length
* 78530 78629: gap of unknown length
* 78630 85476: contig of 6846 bp in length
* 85476 85576: gap of unknown length
* 85576 97151: contig of 11576 bp in length
* 97152 97251: gap of unknown length
* 97252 113111: contig of 15860 bp in length
* 113112 113211: gap of unknown length
* 113212 129870: contig of 16659 bp in length
* 129871 145568: contig of 15598 bp in length
* 145569 145668: gap of unknown length
* 145669 164596: contig of 18928 bp in length
* 164597 164696: gap of unknown length
* 164697 185982: contig of 21286 bp in length.
Location/Qualifiers
1..185982
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/clone="RP11-729A4"
1..1724
/note="assembly_name:Contig3"
1825..2971
/note="assembly_name:Contig4"
3072..4660
/note="assembly_name:Contig5"
4761..5930
/note="assembly_name:Contig6"
clone_end:T7
vector_side:right"
6031..8127
/note="assembly_name:Contig7"
8228..10160
/note="assembly_name:Contig8"
10261..13236
/note="assembly_name:Contig9"
13337..16603
/note="assembly_name:Contig10"
16704..20684
/note="assembly_name:Contig11"
20785..24835
/note="assembly_name:Contig12"
24936..29727
/note="assembly_name:Contig13"
clone_end:SP6
vector_side:right"
29828..34947
/note="assembly_name:Contig14"
35048..41164
/note="assembly_name:Contig15"
41265..47050
/note="assembly_name:Contig16"
47151..53084
/note="assembly_name:Contig17"
```



misc\_feature 53185. .60624  
misc\_feature 50725. .69057  
misc\_feature 69158. .78529  
misc\_feature 78630. .85475  
misc\_feature 85576. .97151  
misc\_feature 97252. .113111  
misc\_feature 113212. .129870  
misc\_feature 129971. .145568  
misc\_feature 145669. .164596  
misc\_feature 164697. .185982  
misc\_feature 185982. .21286

BASE COUNT 53704 a 39120 c 39561 g 51192 t 2405 others  
ORIGIN

Query Match 64.4%; Score 483.2; DB 2; Length 185982;  
Best Local Similarity 99.0%; Pred. No. 3.8e-105;  
Matches 507; Conservative 0; Mismatches 3; Indels 2; Gaps 2;  
QY 241 ATCTAGCAGCTGCTCGGTTAGCAGAA-GGGGACTCTGTATCTGCTCTGAGCTTC 299  
Db 24936 ATCTAGCAGCTGCTCGGTTAGCAGAAATGGGACTCTGTATCTGCTCTGAGCTTC 24995  
QY 300 TTCAGCTGATTATTAATGGAACACAGAGTAGATATTGTTGCAATTAAGTGAATATTA 359  
Db 24996 TTCAGCTGATTATTAATGGAACACAGAGTAGATATTGTTGCAATTAAGTGAATATTA 25055  
QY 360 TGAGAATCATCATAGCAAACTTCACAGTTGATCAAGGATCTGCTTCAATCTGGCC 419  
Db 25056 TGAGAATCATCATAGCAAACTTCACAGTTGATCAAGGATCTGCTTCAATCTGGCC 25115  
QY 420 -AACTGATGCTTAAAGCAGCTGCAAGACTTCAGAGCTGACAAAAGCAAACTCCAG 478  
Db 25116 AAAGTGAATGTTAAAGCAGCTGCAAGACTTCAGAGCTGACAAAAGCAAACTCCAG 25175  
QY 479 ACTTATTTCTGGAATCTGTTTGTGAGACACTGGCCCATGAAATGCTCTCCAGAAAT 538  
Db 25176 ACTTATTTCTGGAATCTGTTTGTGAGACACTGGCCCATGAAATGCTCTCCAGAAAT 25235  
QY 539 AGTCGGATTTGCTGCAAAATTAATTTGGGCAATTTACAGACATGTGCTTTTTCAGAG 598  
Db 25236 AGTCGGATTTGCTGCAAAATTAATTTGGGCAATTTACAGACATGTGCTTTTTCAGAG 25295  
QY 599 ATTTATTTTAAATTAATTAATTTAAATTAATTAATTAATTAATTAATTAATTAATTA 658  
Db 25296 ATTTATTTTAAATTAATTAATTTAAATTAATTAATTAATTAATTAATTAATTAATTA 25355  
QY 659 GTGTGCAATCTGCTACATGCTGCTAATGAGTTTACATATATTTTACACCCCATGTAAT 718  
Db 25356 GTGTGCAATCTGCTACATGCTGCTAATGAGTTTACATATATTTTACACCCCATGTAAT 25415  
QY 719 CACCACCGAGATCTAGATAAGATGATTCGTT 750  
Db 25416 CACCACCGAGATCTAGATAAGATGATTCGTT 25447

RESULT 5  
AC073231/c  
LOCUS  
DEFINITION Homo sapiens chromosome 17 clone RP11-729A4, WORKING DRAFT  
SEQUENCE, 25 unordered pieces.  
AC073231  
AC073231.1 GI:8440043  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens.

ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 185982)  
Waterston,R.H.  
The sequence of Homo sapiens clone  
Unpublished  
2 (bases 1 to 185982)  
Waterston,R.H.  
Direct Submission  
Submitted (10-JUN-2000) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA

COMMENT

----- Genome Center -----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 25 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1724: contig of 1724 bp in length  
\* 1725 1824: gap of unknown length  
\* 1825 2971: contig of 1147 bp in length  
\* 2972 3071: gap of unknown length  
\* 3072 4660: contig of 1589 bp in length  
\* 4661 4760: gap of unknown length  
\* 4761 5930: contig of 1170 bp in length  
\* 5931 8127: contig of 2097 bp in length  
\* 8128 8227: gap of unknown length  
\* 8228 10160: contig of 1933 bp in length  
\* 10161 10260: gap of unknown length  
\* 10261 13236: contig of 2976 bp in length  
\* 13237 13336: gap of unknown length  
\* 13337 16603: contig of 3267 bp in length  
\* 16604 16703: gap of unknown length  
\* 16704 20684: contig of 3981 bp in length  
\* 20685 20784: gap of unknown length  
\* 20785 24835: contig of 4051 bp in length  
\* 24836 24935: gap of unknown length  
\* 24936 29277: contig of 4792 bp in length  
\* 29278 29827: gap of unknown length  
\* 29828 34947: contig of 5120 bp in length  
\* 34948 35048: gap of unknown length  
\* 35049 41164: contig of 6117 bp in length  
\* 41165 41264: gap of unknown length  
\* 41265 47050: contig of 5786 bp in length  
\* 47051 47151: gap of unknown length  
\* 47152 53084: contig of 5934 bp in length  
\* 53085 53184: gap of unknown length  
\* 53185 60624: contig of 7440 bp in length  
\* 60625 60724: gap of unknown length  
\* 60725 69057: contig of 8333 bp in length  
\* 69058 69157: gap of unknown length  
\* 69158 78529: contig of 9372 bp in length  
\* 78530 78629: gap of unknown length  
\* 78630 85475: contig of 6846 bp in length  
\* 85476 85575: gap of unknown length  
\* 85576 97151: contig of 11576 bp in length  
\* 97152 97252: gap of unknown length  
\* 97253 113111: contig of 15860 bp in length  
\* 113112 113211: gap of unknown length  
\* 113212 129870: contig of 16659 bp in length  
\* 129871 129971: gap of unknown length  
\* 129972 145668: contig of 15598 bp in length  
\* 145669 145696: gap of unknown length  
\* 145697 164597: contig of 18928 bp in length  
\* 164598 164696: gap of unknown length  
\* 164697 185982: contig of 21286 bp in length.

Location/Qualifiers

FEATURES

```
source
1. .185982
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/clone="RP11-729A4"
1. .1724
/note="assembly_name:Contig3"
1825. .2971
/note="assembly_name:Contig4"
3072. .4660
/note="assembly_name:Contig5"
4761. .5930
/note="assembly_name:Contig6"
clone_end:T7
vector_side:right"
6031. .8127
/note="assembly_name:Contig7"
8228. .10160
/note="assembly_name:Contig8"
10261. .13236
/note="assembly_name:Contig9"
13337. .16603
/note="assembly_name:Contig10"
16704. .20684
/note="assembly_name:Contig11"
20785. .24835
/note="assembly_name:Contig12"
24936. .29727
/note="assembly_name:Contig13"
clone_end:SP6
vector_side:right"
29828. .34947
/note="assembly_name:Contig14"
35048. .41164
/note="assembly_name:Contig15"
41265. .47050
/note="assembly_name:Contig16"
47151. .53084
/note="assembly_name:Contig17"
53185. .60624
/note="assembly_name:Contig18"
60725. .69057
/note="assembly_name:Contig19"
69158. .78529
/note="assembly_name:Contig20"
78630. .85475
/note="assembly_name:Contig21"
85576. .97151
/note="assembly_name:Contig22"
97252. .113111
/note="assembly_name:Contig23"
113212. .129870
/note="assembly_name:Contig24"
129971. .145568
/note="assembly_name:Contig25"
145669. .164596
/note="assembly_name:Contig26"
164697. .185982
/note="assembly_name:Contig27"
53704 a 39120 c 39561 g 51192 t 2405 others
BASE COUNT
ORIGIN

Query Match 18.0%; Score 135; DB 2; Length 185982;
Best Local Similarity 99.3%; Pred. No. 5e-22;
Matches 146; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 CGCCTCTGGCAGGTAGACCTTGAAGCAAACTGACGTTCGTAGGTTGTAGGCGGAAT 60
Db 6185 CGCCTCTGGCAGGTAGACCTTGAAGCAAACTGACGTTCGTAGGTTGTAGGCGGAAT 6126
Qy 61 AATTACTGCTGGCAGTCAGCAGCTTCCCAACCGCTTCGTGAGGCGGAGGTGTTATTCGCC 120
Db 6125 AATTACTGCTGGCAGTCAGCAGCTTCCCAACCGCTTCGTGAGGCGGAGGTGTTATTCGCC 6066
```

```
Qy 121 AGTTTGGCACAGGGCACAGGTGTAGA 147
| | | | | | | | | | | | | | | | | |
Db 6065 AGTTTGG-ACAAGGGCACAGGTGTAGA 6040

RESULT 6
AC007088 22295 bp DNA linear PRI 21-DEC-1999
LOCUS Homo sapiens PAC clone RP6-39H21 from Xq23, complete sequence.
DEFINITION AC007088
ACCESSION AC007088
VERSION AC007088.3 GI:5757556
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 22295)
Sulston,J.E. and Waterston,R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
9847074
2 (bases 1 to 22295)
Holmes,A., Wohldman,P. and Le,T.
The sequence of Homo sapiens PAC clone RP6-39H21
Unpublished
3 (bases 1 to 22295)
Waterston,R.H.
Direct Submission
Submitted (16-MAR-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 22295)
Waterston,R.H.
Direct Submission
Submitted (21-AUG-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 22295)
Waterston,R.
Direct Submission
Submitted (21-DEC-1999) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Aug 21, 1999 this sequence version replaced gi:4454629.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
-----
Center project name: H_DA0039H21
-----

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
This sequence was generated from part of bacterial clone contigs of
human chromosome X, constructed by the chromosome X mapping group
at the Sanger Centre, Wellcome Trust Genome Campus, Hinxton, UK.
Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX/
```

SOURCE INFORMATION:

This clone was derived from human PAC library RPCI-6, prepared by Pieter de Jong and coworkers at Roswell Park Cancer Institute, using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one female donor. For further details, see <http://bacpac.med.buffalo.edu/>. The clone is available from Genome Systems, Inc. (<http://www.genomesystems.com>).

VECTOR: pPAC4

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP1-169K13, 200 bp overlap. Actual start of this clone is at base position 1191 of RP1-169K13; actual end is at base position 22295 of RP6-39H21.

FEATURES

source

1. .22295  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="X"  
/map="Xg23"  
/clone="RP6-39H21"  
/clone\_lib="RPCI-6"  
140. .277  
/rpt\_family="MIR"  
451. .568  
/rpt\_family="Alu"  
569. .592  
/rpt\_family="(CA)n"  
941. .988  
/rpt\_family="L2"  
1111. .1417  
/rpt\_family="Alu"  
1791. .1997  
/rpt\_family="L1"  
2728. .2742  
/rpt\_family="(GGGA)n"  
2743. .3033  
/rpt\_family="Alu"  
3034. .3058  
/rpt\_family="(GGGA)n"  
3059. .3432  
/rpt\_family="L2"  
3444. .3953  
/rpt\_family="L2"  
4437. .4548  
/rpt\_family="Alu"  
4578. .4702  
/rpt\_family="Alu"  
4808. .4891  
/rpt\_family="L2"  
5109. .5190  
/rpt\_family="MIR"  
6094. .6113  
/rpt\_family="AT-rich"  
6114. .6411  
/rpt\_family="Alu"  
complement(6184. .6446)  
/db\_xref="GI:2624097"  
6412. .6440  
/rpt\_family="AT-rich"  
6755. .6965  
/rpt\_family="MIR"  
7091. .7279  
/rpt\_family="MIR"  
7355. .7761  
/rpt\_family="MER4-group"  
8324. .8533  
/rpt\_family="MIR"  
9152. .9535  
/rpt\_family="Retroviral"  
9549. .9837  
/rpt\_family="L2"  
10088. .10132

repeat\_region /rpt\_family="(GAAA)n"  
10316. .10407  
/rpt\_family="MIR"  
repeat\_region 10477. .10516  
/rpt\_family="A-rich"  
repeat\_region 11043. .11158  
/rpt\_family="MIR"  
repeat\_region 11926. .11976  
/rpt\_family="L2"  
repeat\_region 12046. .12292  
/rpt\_family="Alu"  
repeat\_region 12293. .12354  
/rpt\_family="polypurine"  
repeat\_region 13227. .13509  
/rpt\_family="L1"  
repeat\_region 13510. .14211  
/rpt\_family="L1"  
repeat\_region 14212. .14301  
/rpt\_family="L1"  
repeat\_region 14307. .14390  
/rpt\_family="Alu"  
repeat\_region 14391. .14682  
/rpt\_family="L1"  
repeat\_region 14883. .14942  
/rpt\_family="(CA)n"  
repeat\_region 15338. .15362  
/rpt\_family="(TAAA)n"  
repeat\_region 15363. .15644  
/rpt\_family="Alu"  
repeat\_region 15719. .16346  
/rpt\_family="L1"  
repeat\_region 16347. .16647  
/rpt\_family="Alu"  
repeat\_region 16648. .16895  
/rpt\_family="L1"  
repeat\_region 16930. .17081  
/rpt\_family="L1"  
repeat\_region 18165. .18256  
/rpt\_family="L1"  
repeat\_region 18418. .18545  
/rpt\_family="L1"  
repeat\_region 19068. .19081  
/rpt\_family="purine-rich"  
repeat\_region 19082. .19390  
/rpt\_family="Alu"  
repeat\_region 19391. .19450  
/rpt\_family="purine-rich"  
repeat\_region 20741. .21429  
/rpt\_family="L1"  
repeat\_region 21430. .21719  
/rpt\_family="Alu"  
repeat\_region 21720. .21790  
/rpt\_family="L1"  
repeat\_region 21949. .22042  
/rpt\_family="L1"  
repeat\_region 22043. .22084  
/rpt\_family="(CA)n"  
repeat\_region 22065. .22150  
/rpt\_family="L1"  
repeat\_region 22236. .22295  
/rpt\_family="Alu"

BASE COUNT 7309 a 4342 c 4516 g 6128 t

Query Match 8.2%; Score 61.6; DB 9; Length 22295;

Best Local Similarity 63.5%; Pred. No. 0.0002;

Matches 94; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Qy 587 TCTTTTTCAGACATTATTTTAACTATTTAAAAATATTAAACATGGTACAATTT 646

Db 1759 TTTTCTGCCATATATATTTTCCCTCATTTCTTTTAAAGTTTATGGGATATAA 1818

Qy 647 GCATATAGTGAAGTGGCAATCTTCGCTACATGGCTCAATGAGTTTTTACATATATTC 706

Db	1819	ACATACAGTAAAGTCACAAAATCTCAAGTGAATCACTCACTAAATAAAATTTTACTATTGTGTA	1878
Qy	707	CACCATGTATTAATCACACCGAGATCTAG	734
Db	1879	TACCATGTAAACCACTATCCACATCAAG	1906
RESULT 7			
AC015553/c			
LOCUS			
DEFINITION		AC015553 Homo sapiens 9 BAC RP11-100N10 (Roswell Park Cancer Institute Human BAC Library) complete sequence.	
ACCESSION		AC015553	
VERSION		1	
KEYWORDS		HTG.	
SOURCE		AC015553.21 GI:13489133	
ORGANISM		Homo sapiens.	
		Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1 (bases 1 to 162063)			
REFERENCE			
AUTHORS			
		Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbarta,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Bricket,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,X., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Haviak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louisghe,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J.J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokwako,S., Ogum,H., Okunoye,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D. and Gibbs,R.	
TITLE		Direct Submission	
JOURNAL		Unpublished	
REFERENCE		2 (bases 1 to 162063)	
AUTHORS		Worley,K.C.	
TITLE		Direct Submission	
JOURNAL		Submitted (17-NOV-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	
REFERENCE		3 (bases 1 to 162063)	
AUTHORS		Worley,K.C.	
TITLE		Direct Submission	
JOURNAL		Submitted (30-MAR-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	

```
98412 acgttaactta(n)ncaagggtccc
98413 cgttaacttan(n)caagggtccc
113818 tgatactgaa(n)cnctataaaaa
113821 tactgaancc(n)ctataaaaaa
113849 aaaaaaaa(n)gactaaaaata
122621 atggaccatg(n)tgctttgtga
133372 gttataaaa(n)caatacaacta
134847 tctttactg(n)ggccctatta
160961 aaaaaaaat(c)caaaaacttg

----- Distribution of Quality < 40 Bases -----
# bases
2000 5 10 15 20 25 30 35 40
1500 * * * * *
1000 * * * * *
500 * * * * *
0 0 0 0 0 0 0 0 0

Version: 1.01 qxfo.
Location/Qualifiers
1. 162063
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-100N10"
706..734
/rpt_family="AT-rich"
complement(738..930)
/rpt_family="L1MA4"
complement(943..1086)
/rpt_family="L1ME3A"
1979..2255
/rpt_family="L1MA8"
2560..2722
/rpt_family="AluJb"
2723..2831
/rpt_family="GA-rich"
2832..2859
/rpt_family="(GAAA)n"
2871..3150
/rpt_family="AluY"
3151..3335
/rpt_family="(GAAA)n"
3336..3346
/rpt_family="AluY"
3347..3481
/rpt_family="L1M4"
3482..4108
/rpt_family="L1MA4"
4109..4207
/rpt_family="L1M4"

Query Match 7.7%; Score 57.8; DB 9; Length 162063;
Best Local Similarity 59.5%; Pred. NO. 0.0013;
Matches 116; Conservative 0; Mismatches 77; Indels 2; Gaps 1;

QY 556 AATAAATTGGCAATCTACAGACATGTCTTTTCAGAGATTATTTTAAATAA 615
Db 68440 AATAATGTCAGGAGTCACATACAACTGTCTTATCCAGATGAAGATTATATT 68381
QY 616 CTTA--TTTAAATATTAACATGGTACAAATTTCATATAGTGAAGTGTGCAATCTTCG 673
```

```
Db 68380 ATTAGCTTTTGTACTTTTTTTTACTAGGTCACATATACACAGTAAAGTCACATCTTAA 68321
QY 674 CTACATGGCTCAATGAGCTTTTACATATATTCACCCATGTAATCACACCGAGATCTA 733
Db 68320 TTCCACAGCTCGATGAATTTTCTATGTATACAAATGGAACACGACCTAGATGAA 68261
QY 734 GAATAGAATGATTTC 748
Db 68260 AATAAGAACAAATTC 68246

RESULT 8
AL157884/c
LOCUS
DEFINITION
Human DNA sequence from clone RP11-462B18 on chromosome
9p13.1-21.1, complete sequence.
ACCESSION
AL157884
VERSION
AL157884.9 GI:15029097
KEYWORDS
HTG.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 176932)
AUTHORS
Whitehead,S.
TITLE
Direct Submission
JOURNAL
Submitted (13-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
COMMENT
On Jul 29, 2001 this sequence version replaced gi:15026899.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em:, EMBL; Sw:,
SWISSPROT; Tr:, TrEMBL; Wp:, WormPEP; Information on the WormPEP
database can be found at
http://www.sanger.ac.uk/Projects/C-elegans/WormPEP This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9
RP11-462B18 is from the library RPCI-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pSACE3.6
This sequence is the entire insert of clone RP11-462B18 The true
right end of clone RP11-555J4 is at 30638 in this sequence. The
true right end of clone RP11-1K21 is at 97783 in this sequence.

FEATURES
Location/Qualifiers
1..176932
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/map="p13.1-21.1"
/clone="RP11-462B18"
/clone_lib="RPCI-11.2"
76326..77014
/note="CpG island"
/misc_feature
/evidence=not_experimental

BASE COUNT 54421 a 37529 c 35943 g 49039 t
ORIGIN
```

```

Query Match          7.7%; Score 57.8; DB 9; Length 176932;
Best Local Similarity 59.5%; Pred. No. 0.0013;
Matches 116; Conservative 0; Mismatches 77; Indels 2; Gaps 1;

Qy 556 AATAAATTTGGCAATCTCAGACAAATGTCGCTTTTTCAGAGATTTATTTTAAATAAA 615
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 128943 AATAATGTCAGGCAGTCACACATACATCTCTTTTGATCCAAAGATTGAAGATTTATTT 128884
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 616 CTTA--TTTAAAAATATTAAATGGTACAAATTTGCATATAGTGAAGTGCCTCAATCTTCG 673
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 128883 ATTAGCTTTGTACTTTTTTTACTGAGGTCAACATATACACAGTAAAGTGCACTCATCTTAA 128824
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 674 CTACATGGCTCAATGAGTTTTTACATATATATCCACCCATGTAATCACCACCGAGATCTA 733
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 128823 TTGCACAGCTCGATGATTTTATCTATGTATACAAACATGCAACCGACCTGATGAA 128764
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 734 GAATAGAATGTATTC 748
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 128763 AATAAGAAACAATTC 128749
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
AP000647 118819 bp DNA linear PRI 30-MAR-2001
LOCUS Homo sapiens genomic DNA, chromosome 11q, clone: CMB9-75I23,
DEFINITION complete sequences.
ACCESSION AP000647
VERSION AP000647.4 GI:13488749
KEYWORDS HTG.
SOURCE Homo sapiens DNA, clone: CMB9-75I23.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Homo sapiens genomic DNA
JOURNAL Published Only in DataBase (1999)
REFERENCE 2 (bases 1 to 118819)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (28-OCT-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Shuhiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@gs.c.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT On Mar 29, 2001 this sequence version replaced gi:11320561.
FEATURES
     location/Qualifiers
     source
       1..118819
       /organism="Homo sapiens"
       /db_xref="taxon:9606"
       /chromosome="11"
       /map="11q"
       /clone="CMB9-75I23"

BASE COUNT 33872 a 23591 c 22856 g 38500 t
ORIGIN

Query Match          7.5%; Score 56.6; DB 9; Length 118819;
Best Local Similarity 60.9%; Pred. No. 0.0027;
Matches 92; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Qy 600 TTTATTTTAACTTAACTTTATTTAAAAATATTAAACATGGTGACAAATTCGATATAGTGAAG 659
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 36871 TTTCGCTGTAATTCCTTTTCTTTATATATTTTGTGAGATACTATTGCATATAGTGAAG 36930
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 660 TGTGAAATCTTCGCTACATGGCTCAATGAGTTTTTACATATATTTCCACCCATGTAAATC 719
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 36931 TACACAAGCTTTAACTACACTGTTTCGGTGAAGTTTGTGACAAATGTATAAGCCCAATGTATCC 36990
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 720 ACCACCGAGATCTAGAAATGTATTCCT 750
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

134643 138972 contig of 4330 bp in length  
139073 143377 contig of 4265 bp in length  
143438 146720 contig of 3283 bp in length  
146821 149011 contig of 2191 bp in length  
149112 151132 contig of 2021 bp in length  
151233 154403 contig of 3171 bp in length  
154504 156567 contig of 2064 bp in length  
156668 158701 contig of 2034 bp in length  
158802 160732 contig of 1931 bp in length  
160833 162522 contig of 1690 bp in length  
162623 164124 contig of 1502 bp in length  
164225 165574 contig of 1350 bp in length  
165675 167245 contig of 1571 bp in length  
167346 168587 contig of 1242 bp in length  
168688 169757 contig of 1070 bp in length  
169858 170997 contig of 1140 bp in length  
171098 172148 contig of 1051 bp in length  
172249 173359 contig of 1111 bp in length

Sequence updated (26-May-2000).

\* NOTE: This is a 'working draft' sequence. It currently consists of 35 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. \* This record will be updated with the finished sequence. \* as soon as it is available and the accession number will be preserved.

1 15727: contig of 15727 bp in length  
\* 15728 15827: gap of 100 bp  
\* 15828 27186: contig of 11359 bp in length  
\* 27187 27286: gap of 100 bp  
\* 27287 37387: contig of 10101 bp in length  
\* 37388 37487: gap of 100 bp  
\* 37488 47288: contig of 9801 bp in length  
\* 47289 47388: gap of 100 bp  
\* 47389 55992: contig of 8504 bp in length  
\* 55993 56092: gap of 100 bp  
\* 56093 63632: contig of 7540 bp in length  
\* 63633 63732: gap of 100 bp  
\* 63733 73241: contig of 9509 bp in length  
\* 73242 73341: gap of 100 bp  
\* 73342 80659: contig of 7318 bp in length  
\* 80660 80759: gap of 100 bp  
\* 80760 89061: contig of 8302 bp in length  
\* 89062 89161: gap of 100 bp  
\* 89162 94975: contig of 5814 bp in length  
\* 94976 95075: gap of 100 bp  
\* 95076 100459: contig of 5384 bp in length  
\* 100460 100559: gap of 100 bp  
\* 100560 106303: contig of 5744 bp in length  
\* 106304 106403: gap of 100 bp  
\* 106404 111824: contig of 5421 bp in length  
\* 111825 111924: gap of 100 bp  
\* 111925 117598: contig of 5674 bp in length  
\* 117599 117698: gap of 100 bp  
\* 117699 123774: contig of 6076 bp in length  
\* 123775 123874: gap of 100 bp  
\* 123875 128021: contig of 4147 bp in length  
\* 128022 128121: gap of 100 bp  
\* 128122 134542: contig of 6421 bp in length  
\* 134543 134642: gap of 100 bp  
\* 134643 138972: contig of 4330 bp in length  
\* 138973 139072: gap of 100 bp  
\* 139073 143337: contig of 4265 bp in length  
\* 143338 143437: gap of 100 bp  
\* 143438 146720: contig of 3283 bp in length  
\* 146721 146820: gap of 100 bp  
\* 146821 149011: contig of 2191 bp in length  
\* 149012 149111: gap of 100 bp  
\* 149112 151132: contig of 2021 bp in length  
\* 151133 151232: gap of 100 bp  
\* 151233 154403: contig of 3171 bp in length  
\* 154404 154503: gap of 100 bp

\* 154504 156567: contig of 2064 bp in length  
\* 156568 156667: gap of 100 bp  
\* 156668 158701: contig of 2034 bp in length  
\* 158702 158801: gap of 100 bp  
\* 158802 160732: contig of 1931 bp in length  
\* 160733 160832: gap of 100 bp  
\* 160833 162522: contig of 1690 bp in length  
\* 162523 162622: gap of 100 bp  
\* 162623 164124: contig of 1502 bp in length  
\* 164125 164224: gap of 100 bp  
\* 164225 165574: contig of 1350 bp in length  
\* 165575 165674: gap of 100 bp  
\* 165675 167245: contig of 1571 bp in length  
\* 167246 167345: gap of 100 bp  
\* 167346 168587: contig of 1242 bp in length  
\* 168588 168687: gap of 100 bp  
\* 168688 169757: contig of 1070 bp in length  
\* 169758 169857: gap of 100 bp  
\* 169858 170997: contig of 1140 bp in length  
\* 170998 171097: gap of 100 bp  
\* 171098 172148: contig of 1051 bp in length  
\* 172149 172248: gap of 100 bp  
\* 172249 173359: contig of 1111 bp in length.

FEATURES

source

Location/Qualifiers  
1..173359  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11q22"  
/clone="RP11-750P5"  
1..15727  
/note="assembly\_fragment"  
15828..27186  
/note="assembly\_fragment clone\_end:SP6 vector\_side:right"  
27287..37387  
/note="assembly\_fragment"  
37488..47288  
/note="assembly\_fragment"  
47389..55992  
/note="assembly\_fragment"  
56093..63632  
/note="assembly\_fragment"  
63733..73241  
/note="assembly\_fragment"  
73342..80659  
/note="assembly\_fragment clone\_end:T7 vector\_side:right"  
80760..89061  
/note="assembly\_fragment"  
89162..94975  
/note="assembly\_fragment"  
95076..100459  
/note="assembly\_fragment"  
100560..106303  
/note="assembly\_fragment"  
106404..111824  
/note="assembly\_fragment"  
111925..117598  
/note="assembly\_fragment"  
117699..123774

Query Match 7.5%; Score 56.6; DB 2; Length 173359;  
Best Local Similarity 60.9%; Pred.No. 0.0026;  
Matches 92; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Qy 600 TTTATTTTAAATAACTTATTTAAAAATTAACATGGTACAATTTGCATATAGTGAAG 659  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 130615 TTTGCTGTTAAATTCCTTTTCTTTATATATTTTGAGATACTATTGACATATAGTGAAG 130556  
Qy 650 TGTGCAAAATCTTCGTACATGCTCAATGAGTTTTTTACATATTTTCCACCCATGTATTC 719  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 130555 TACACAAGTCTTAACACACACTGTTGCGTGAGTTTGGACAAATGTATGACCAATGTATCC 130496  
Qy 720 ACCACCGAGATCTAGNAATAGATGTATTCGT 750

DB 130495 CATACCTCATGAGGAGAGCAATTTCAT 130465

RESULT 11

AC114686

LOCUS

DEFINITION

AC114686

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AC114686 68444 bp DNA linear HTG 10-APR-2002  
Homo sapiens chromosome 17 clone RP11-1D5 map 17, LOW-PASS SEQUENCE  
SAMPLING.  
AC114686  
AC114686.2 GI:20128626  
HTG; HTGS\_PHASED.  
Homo sapiens.  
Homo sapiens.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 68444)  
Birren, B., Linton, L., Nusbaum, C. and Lander, E.

Homo sapiens chromosome 17, clone RP11-1D5  
Unpublished

2 (bases 1 to 68444)

Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,

Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,

Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,

Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,

Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,

Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,

Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,

Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,

Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,

Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,

Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,

McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,

Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,

Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,

Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,

Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,

Roy, A., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,

Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,

Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,

Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,

Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,

Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (10-MAR-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 68444)

Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,

Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,

Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,

Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,

Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,

Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,

Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,

Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,

Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,

AC114686

LOCUS

DEFINITION

AC114686

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AC114686 68444 bp DNA linear HTG 10-APR-2002  
Homo sapiens chromosome 17 clone RP11-1D5 map 17, LOW-PASS SEQUENCE  
SAMPLING.  
AC114686  
AC114686.2 GI:20128626  
HTG; HTGS\_PHASED.  
Homo sapiens.  
Homo sapiens.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 68444)  
Birren, B., Linton, L., Nusbaum, C. and Lander, E.

Homo sapiens chromosome 17, clone RP11-1D5  
Unpublished

2 (bases 1 to 68444)

Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,

Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,

Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,

Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,

Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,

Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,

Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,

Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,

Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,

Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,

Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,

McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,

Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,

Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,

Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,

Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,

Roy, A., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,

Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,

Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,

Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,

Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,

Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Apr 10, 2002 this sequence version replaced gi:19311218.

All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L26108  
Center clone name: L\_D\_5  
-----

\* NOTE: This record contains 83 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

\* 1 712: contig of 712 bp in length  
\* 713 812: gap of 100 bp  
\* 813 1552: contig of 740 bp in length  
\* 1553 1652: gap of 100 bp  
\* 1653 2381: contig of 729 bp in length  
\* 2382 2481: gap of 100 bp  
\* 2482 3209: contig of 728 bp in length  
\* 3210 3309: gap of 100 bp  
\* 3310 4047: contig of 738 bp in length  
\* 4048 4147: gap of 100 bp  
\* 4148 4852: contig of 705 bp in length  
\* 4853 4952: gap of 100 bp  
\* 4953 5690: contig of 738 bp in length  
\* 5691 5790: gap of 100 bp  
\* 5791 6501: contig of 711 bp in length  
\* 6502 6601: gap of 100 bp  
\* 6602 7330: contig of 729 bp in length  
\* 7331 7430: gap of 100 bp  
\* 7431 8147: contig of 717 bp in length  
\* 8148 8247: gap of 100 bp  
\* 8248 8962: contig of 715 bp in length  
\* 8963 9062: gap of 100 bp  
\* 9063 9785: contig of 723 bp in length  
\* 9786 9885: gap of 100 bp  
\* 9886 10619: contig of 734 bp in length  
\* 10620 10719: gap of 100 bp  
\* 10720 11417: contig of 698 bp in length  
\* 11418 11517: gap of 100 bp  
\* 11518 12337: contig of 720 bp in length  
\* 12338 12337: gap of 100 bp  
\* 12338 13069: contig of 732 bp in length  
\* 13070 13169: gap of 100 bp  
\* 13170 13894: contig of 725 bp in length  
\* 13895 13994: gap of 100 bp  
\* 13995 14720: contig of 726 bp in length  
\* 14721 14820: gap of 100 bp  
\* 14821 15548: contig of 728 bp in length  
\* 15549 15648: gap of 100 bp  
\* 15649 16363: contig of 715 bp in length  
\* 16364 16463: gap of 100 bp  
\* 16464 17180: contig of 717 bp in length  
\* 17181 17280: gap of 100 bp  
\* 17281 18014: contig of 734 bp in length  
\* 18015 18114: gap of 100 bp  
\* 18115 18842: contig of 728 bp in length  
\* 18843 18942: gap of 100 bp  
\* 18943 19668: contig of 726 bp in length  
\* 19669 19768: gap of 100 bp  
\* 19769 20497: contig of 729 bp in length  
\* 20498 20597: gap of 100 bp







for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA  
98105, USA  
On Mar 27, 2001 this sequence version replaced gi:8247780.

COMMENT On Mar 27, 2001 this sequence replaced gi:8247780.

----- Genome Center  
Center: Multimerase Sequencing Center  
Center code: UTMSC  
Web site: [http://chroma.mbt.washington.edu/msg\\_www](http://chroma.mbt.washington.edu/msg_www)  
Contact: [leerowen@systemsbiology.org](mailto:leerowen@systemsbiology.org)

```
----- Summary Statistics
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 90% of reads
Chemistry: Dye-primer Big Dye; 10% of reads
Assembly program: Phrap; version 0.990399
```

```

FEATURES
    source
        Location/Qualifiers
            1..190565
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="14"
                /map="14q31"
                /clone="RP11-325L17"
                /clone_lib="RPC1 human BAC library 11"

```

unsure	985. .1035	/note="low quality data"
unsure	3865. .3870	/note="low quality data"
unsure	14875. .14880	/note="low quality data"
unsure	17060. .17095	/note="low quality data"
unsure	17488. .17535	/note="low quality data"
unsure	19433	/note="low quality data"
unsure	46474. .46580	/note="low quality data"
unsure	46695. .46700	/note="low quality data"
unsure	55163. .55242	/note="low quality data"
unsure	110860. .110875	/note="low quality data"
unsure	113613. .113617	/note="low quality data"
misc_feature	116230. .116782	/note="Single clone coverage. PCR and
unsure	117677. .117880	/note="low quality data"
unsure	117730. .117748	/note="low quality data"
unsure	118505. .118736	/note="low quality data"
unsure	144847. .144850	/note="low quality data"
unsure	152067. .152071	/note="low quality data"
unsure	158953. .158990	/note="low quality data"
unsure	177622. .177705	/note="low quality data"
BASE COUNT	50184 a 45565 c 44821 g 49995 t	/note="low quality data"

Query Match	7.4%	Score 55.2	DB 9	Length 190565
Best Local Similarity	62.1%	Pred. No. 0.0055		
Matches 82	Conservative	0	Mismatches 50	Indels 0
				Gaps 0

QY	617	TTATTTAAAAATATTACATCGGTACAAATTTCATATAGTGAAGTGTGCAAAATCTTCGGTA	676
Db	27551	TTATACTTATTTTGGTTAAAGTATATTACATACATAAAGTGTGACACCTTAAGTG	27492
QY	677	CATGGCTCAATGAGTTTTTACATATATTTCCACCCATGTAAATCACCCCGAGACTAGAA	736

Db	27491	TACACTGGATGAATTTTACATAAGTATACACCTTGCAACCAACCAAGATTAAAGC	27432
Qy	737	TAGAATGTATTC	748
Db	27431	ACAAACAGTTC	27420
RESULT 15			
AC114778			
LOCUS	AC114778	207558 bp	DNA linear HTG 08-AUG-2002
DEFINITION	Homo sapiens chromosome 2 clone RP11-51605, WORKING DRAFT SEQUENCE, 5 unordered pieces.		
ACCESSION	AC114778		
VERSION	AC114778.4	GI:2138707	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 207558)		
AUTHORS	Waterston,R.H.		
TITLE	The sequence of Homo sapiens clone		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 207558)		
AUTHORS	Waterston,R.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-MAR-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA		
REFERENCE	3 (bases 1 to 207558)		
AUTHORS	Waterston,R.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (08-AUG-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA		
COMMENT	On Aug 8, 2002 this sequence version replaced gi:21541936.		

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/Index.shtml
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: H_NH0516005
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 205660 bases at least Q40
Consensus quality: 206259 bases at least Q30
Consensus quality: 206519 bases at least Q20
Insert size: 197000; agarose-fp
Insert size: 207158; sum-of-contigs
Quality coverage: 0.00 in Q20 bases; agarose-fp
Quality coverage: 7.74 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 17064: contig of 17064 bp in length
* 17065 17164: gap of unknown length
* 17165 60411: contig of 43247 bp in length
* 60412 60511: gap of unknown length
* 60512 91942: contig of 31431 bp in length

```

91943 92042: gap of unknown length  
92043 206281: contig of 114239 bp in length  
206282 206381: gap of unknown length  
206382 207558: contig of 1177 bp in length.

## FEATURES

```

source      1. .207558
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="2"
            /clone="RP11-51605"
            1. .17064
            /note="assembly_name:Contig12"
            clone_end:SP6
            vector_side:right
            17165. .60411
            /note="assembly_name:Contig14"
            clone_end:T7
            vector_side:left
            60512. .91942
            /note="assembly_name:Contig13"
            92043. .206281
            /note="assembly_name:Contig15"
            206382. .207558
            /note="assembly_name:Contig5"
COUNT      63196 a 41834 c 33615 g 62513 t

```

```

Query Match.          7.3%; Score 55; DB 2; Length 207558;
Best Local Similarity 68.5%; Pred. No. 0.0061;
Matches 76; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 638 GTACAATTTCATATAGTGAAGTGTGCAATCTTCGGTACATGGCTCAATGAGTTTTTAC 697
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 110423 GTGTAATTTACATACAGGAAAGTAAACATATTTTAAATTATATAGCTCAACAAATTTTAC 110482

QY 698 ATATATTTCCACCCTGTATTCACCCGAGATCTAGAAATAGATGTTATTC 748
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 110483 ATATGTACACCCCATGTACCAACCAACCCGAGATCAAGATATAAAACACTTC 110533

```

Search completed: April 11, 2003, 20:44:46  
Job time : 2556.65 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 14:50:38 ; Search time 124.63 Seconds  
(without alignments)  
13552.085 Million cell updates/sec

Title: US-09-914-152-3\_COPY\_1\_750

Perfect score: 750

Sequence: 1 cgcctctgcaaggtagacc.....ctagaatagaatgtattcgt 750

Scoring table: IDENTITY\_NUC

Gapex 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	750	100.0	10562	21	Human beta3Gal-T5
2	121	16.1	807	20	Human gene express
3	53.6	7.1	73334	24	Chemically treated
4	53.6	7.1	73334	24	Human immune syste
5	52.4	7.0	39887	22	Human immune/haema
6	52.4	7.0	39887	22	Human immune/haema
7	46.2	6.2	1341	22	Human immune/haema
8	45.8	6.1	6012	24	Chemically treated
9	45.8	6.1	6012	24	Signal transductio

10	45.6	6.1	809	22	AAL20997
11	45.4	6.1	3619	22	AAS39742
12	45.4	6.1	3619	22	AAK90083
13	45.2	6.0	45546	20	AAK23520
14	44	5.9	600	22	ABA62819
15	44	5.9	600	22	ABA30104
16	44	5.9	600	22	AAK11220
17	44	5.9	600	22	AAK37018
18	44	5.9	600	22	AAI17864
19	44	5.9	600	22	AAI42837
20	43.8	5.8	12293	24	ABN95654
21	43.2	5.8	581	23	ABV51286
22	42.8	5.7	588	22	ABA63119
23	42.8	5.7	588	22	ABA30370
24	42.8	5.7	588	22	AAK11547
25	42.8	5.7	588	22	AAK37319
26	42.8	5.7	588	22	AAI18150
27	42.8	5.7	588	22	AAI43158
28	42.8	5.7	588	24	ABS11312
29	42.6	5.7	459	21	AAK07476
30	42.6	5.7	459	21	AAZ42984
31	42.6	5.7	459	22	AAK46862
32	42.6	5.7	459	24	ABK81640
33	42.6	5.7	25837	22	AAK85952
34	42.6	5.7	50000	20	AAK23517
35	42.6	5.7	147708	23	ABQ88154
36	41.6	5.5	496	23	ABV55578
37	41.4	5.5	290	22	AAK28563
38	41.4	5.5	474	22	AAK02351
39	41.4	5.5	14287	24	ABN80032
40	41.2	5.5	1402	22	AAK36774
41	41.2	5.5	1402	22	AAK36775
42	41.2	5.5	1402	22	AAK36776
43	41.2	5.5	24259	22	AAK46692
44	41	5.5	215	22	AAK59024
45	41	5.5	445	22	AAI91093

ALIGNMENTS

RESULT 1  
AAA93876  
ID AAA93876 standard; DNA; 10562 BP.  
AC AAA93876;  
XX  
DT 15-JAN-2001 (first entry)  
XX Human beta3Gal-T5 encoding DNA.  
DE  
DE Beta-1,3 galactose transferase; treatment; diagnosis; cancer; human;  
KW digestive system; beta3Gal-T5; ds.  
XX  
OS Homo sapiens.  
XX  
PN W0200050608-A1.  
XX  
PD 31-AUG-2000.  
XX  
XX 24-FEB-2000; 2000WO-JP01070.  
XX  
XX 25-FEB-1999; 99JP-0047571.  
XX  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
PA  
PI Narimatsu H, Isshiki S, Togayachi A, Sasaki K;  
XX  
XX WPI; 2000-549409/50.  
XX  
PT Beta-1,3 galactose transferase and DNA encoding it, useful for  
PT synthesis of type 1 sialyl Lewis, a carbohydrate for treatment of  
PT digestive system cancer

XX Claim 31; Page 103-111; 123pp; Japanese.

XX This invention relates to a polypeptide (I) with beta-1,3 galactose transferase activity, or variants of (I) comprising amino acid additions, deletions and/or substitutions. Included in the invention is DNA encoding all or part of (I); expression vectors containing the DNA, host cells transformed by the vectors; a method for the preparation of the polypeptide by culture of the transformants or by expression in the milk of a transgenic mammal, and antibodies recognising (I). The Beta-1,3 galactose transferase protein transfers galactose by beta-1,3 bonding to N-acetylglucosamine present in a non-cyclic carbohydrate chain (such as GLNAbeta1-3galbeta1-4Glc) to give Galbeta1-3GlcNAc. The protein and DNA encoding it are useful for the treatment and diagnosis of cancer of the digestive system. The present sequence represents a Beta3gal-15 encoding DNA sequence.

XX Sequence 10562 BP; 2610 A; 2415 C; 2574 G; 2963 T; 0 other;

Query Match 100.0%; Score 750; DB 21; Length 10562;  
Best Local Similarity 100.0%; Pred. No. 4.5e-193;  
Matches 750; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTCTGCGAGGTAGACCTTGAAGGCAAACTGAGTTTGTAGACGGAAT 60  
Db 1 CCCTCTGCGAGGTAGACCTTGAAGGCAAACTGAGTTTGTAGACGGAAT 60

Qy 61 AATTACTGCTGGCATGACGACTTCCCAACCGTCTGTGAGGAGGAGTATTGCC 120  
Db 61 AATTACTGCTGGCATGACGACTTCCCAACCGTCTGTGAGGAGGAGTATTGCC 120

Qy 121 AGTTTGGCAAGGGCACAGGTGTAGAACACGTAAGTGCCTGGCGGTGTACACCACC 180  
Db 121 AGTTTGGCAAGGGCACAGGTGTAGAACACGTAAGTGCCTGGCGGTGTACACCACC 180

Qy 181 ACTGTTTGTAGCTGAGTGTGAACAGGCGCTTCTGATTCACAAATTCCTCATTCCTTC 240  
Db 181 ACTGTTTGTAGCTGAGTGTGAACAGGCGCTTCTGATTCACAAATTCCTCATTCCTTC 240

Qy 241 ATCCTAGCAGGCTGCTGGGTTAGCAGAGGGGACCTCTGTATCTGCTCTGCAGCTTCT 300  
Db 241 ATCCTAGCAGGCTGCTGGGTTAGCAGAGGGGACCTCTGTATCTGCTCTGCAGCTTCT 300

Qy 301 TCAGCTGATTTTATAATGGAAACACAGATAGATATTGATTGGCAATAGTGAATATTAT 360  
Db 301 TCAGCTGATTTTATAATGGAAACACAGATAGATATTGATTGGCAATAGTGAATATTAT 360

Qy 361 GAGATCATATAGCAAACTTACAGTTTGTATCAAGGATCCTGCTTCAATATCTGGCCA 420  
Db 361 GAGATCATATAGCAAACTTACAGTTTGTATCAAGGATCCTGCTTCAATATCTGGCCA 420

Qy 421 ACTGATGTATAAGCAGCTGCAAGACTTCAGAGCTGACAAAAAGCAAACTCCAGAC 480  
Db 421 ACTGATGTATAAGCAGCTGCAAGACTTCAGAGCTGACAAAAAGCAAACTCCAGAC 480

Qy 481 TTATTTCTCGAATCTGTTTGTAGACACTGCCCATGAAATGCTCTCCAGAAATAG 540  
Db 481 TTATTTCTCGAATCTGTTTGTAGACACTGCCCATGAAATGCTCTCCAGAAATAG 540

Qy 541 TCGGATTTCTGGTCAATAAATTTGGCAATTTACAGACATGCTGCTTTTTCAGAGAT 600  
Db 541 TCGGATTTCTGGTCAATAAATTTGGCAATTTACAGACATGCTGCTTTTTCAGAGAT 600

Qy 601 TTATTTTAAATAAATTTTAAAAATATTAAATGGTACAAATTTGCATATAGTGAAGT 660  
Db 601 TTATTTTAAATAAATTTTAAAAATATTAAATGGTACAAATTTGCATATAGTGAAGT 660

Qy 661 GTGCAAACTTCGCTACATGGCTCAATGAGTTTTCATATATTTCCACCACCATGTAATCA 720  
Db 661 GTGCAAACTTCGCTACATGGCTCAATGAGTTTTCATATATTTCCACCACCATGTAATCA 720

Qy 721 CCACCGAGATCTAGATAGATATTCGT 750  
Db 721 CCACCGAGATCTAGATAGATATTCGT 750

Db 721 CCACCGAGATCTAGATAGATATTCGT 750

RESULT 2

AAZ16242

ID AAZ16242 standard; cDNA; 807 BP.

XX AAZ16242;

XX 12-OCT-1999 (first entry)

DE Human gene expression product cDNA sequence SEQ ID NO:3712.

XX Human; gene; gene expression product; diagnosis; therapy; probe; detection; mapping; tissue typing; profiling; forensic; cancer; genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.

XX Homo sapiens.

XX WO938972-A2.

XX 05-AUG-1999.

XX 28-JAN-1999; 99WO-US01619.

XX 03-APR-1998; 98US-0080666.

PR 28-JAN-1998; 98US-0072910.

PR 24-FEB-1998; 98US-0075954.

PR 31-MAR-1998; 98US-0080114.

PR 03-APR-1998; 98US-0080515.

XX (CHIR ) CHIRON CORP.

PA (HYSE-) HYSEQ INC.

XX Crkvenjakov R, Dickson M, Drmanac R, Drmanac S; Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA; Jones WL, Kassam A, Kennedy GC, Kita D, Labat I; Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C; Stache-Crain B, Sudduth-Klinger J, Williams LT; WPI; 1999-494092/41.

XX Novel human genes and their expression products which are differentially expressed in different cell types

PS Claim 1; Page 1770; 2479pp; English.

The present invention describes a library of human polynucleotides comprising the sequences given in AAZ12532 to AAZ1779. Also described is a method of detecting differentially expressed genes correlated with the cancerous state of a mammalian cell, comprising detecting at least one differentially expressed gene product in a test sample from a cell suspected of being cancerous, where the gene product is encoded by one of the 5248 polynucleotide sequences given in AAZ12532 to AAZ1779. The polynucleotides can be used as a source of primers and probes, which can be used for a variety of purpose, e.g. detection of expression levels, mapping, tissue typing or profiling, forensics, genetic analysis and detection of polymorphisms. Polypeptides encoded by the polynucleotides can be used for raising antibodies for experimental, diagnostic and therapeutic purposes. The polynucleotides may also be used to construct arrays for diagnostics (which may be used to determine function of an encoded protein); and to detect differences in expression levels between two cells (e.g. to identify abnormal or diseased tissue in a human, to identify a genetic predisposition or susceptibility to a disease such as cancer). The polynucleotides of the invention are especially used in the diagnosis, prognosis and management of colorectal cancer, breast cancer, and lung cancer. The polynucleotides can also be used to screen for peptide analogues and antagonists.

Sequence 807 BP; 188 A; 187 C; 190 G; 203 T; 39 other;

Query Match 16.1%; Score 121; DB 20; Length 807;  
Best Local Similarity 79.0%; Pred. No. 7.9e-23;



PR	14-AUG-2000;	2000US-02257515
PR	14-AUG-2000;	2000US-02257518
PR	14-AUG-2000;	2000US-02257529
PR	14-AUG-2000;	2000US-02257538
PR	18-AUG-2000;	2000US-02256779
PR	22-AUG-2000;	2000US-02266681
PR	22-AUG-2000;	2000US-02266688
PR	22-AUG-2000;	2000US-02266868
PR	22-AUG-2000;	2000US-02271882
PR	23-AUG-2000;	2000US-02277009
PR	30-SEP-2000;	2000US-02289824
PR	01-SEP-2000;	2000US-02292887
PR	01-SEP-2000;	2000US-02293433
PR	01-SEP-2000;	2000US-02293444
PR	01-SEP-2000;	2000US-02293454
PR	05-SEP-2000;	2000US-02295909
PR	05-SEP-2000;	2000US-02295913
PR	06-SEP-2000;	2000US-02304537
PR	06-SEP-2000;	2000US-02304338
PR	08-SEP-2000;	2000US-02304348
PR	08-SEP-2000;	2000US-02312422
PR	08-SEP-2000;	2000US-02312424
PR	08-SEP-2000;	2000US-02312444
PR	08-SEP-2000;	2000US-02314133
PR	08-SEP-2000;	2000US-02314144
PR	08-SEP-2000;	2000US-02320800
PR	12-SEP-2000;	2000US-02320811
PR	12-SEP-2000;	2000US-02321968
PR	14-SEP-2000;	2000US-02323937
PR	14-SEP-2000;	2000US-02323998
PR	14-SEP-2000;	2000US-02323998
PR	14-SEP-2000;	2000US-02323999
PR	14-SEP-2000;	2000US-02324001
PR	14-SEP-2000;	2000US-02324001
PR	14-SEP-2000;	2000US-02324011
PR	14-SEP-2000;	2000US-02324033
PR	14-SEP-2000;	2000US-02323064
PR	14-SEP-2000;	2000US-02323065
PR	21-SEP-2000;	2000US-02323065
PR	21-SEP-2000;	2000US-02342724
PR	21-SEP-2000;	2000US-02342724
PR	25-SEP-2000;	2000US-02349977
PR	25-SEP-2000;	2000US-02349988
PR	26-SEP-2000;	2000US-02354844
PR	27-SEP-2000;	2000US-02354834
PR	27-SEP-2000;	2000US-02355836
PR	29-SEP-2000;	2000US-02363627
PR	29-SEP-2000;	2000US-02363627
PR	29-SEP-2000;	2000US-02363668
PR	29-SEP-2000;	2000US-02363669
PR	29-SEP-2000;	2000US-02363700
PR	02-OCT-2000;	2000US-02363700
PR	02-OCT-2000;	2000US-02368602
PR	02-OCT-2000;	2000US-02370737
PR	02-OCT-2000;	2000US-02370738
PR	02-OCT-2000;	2000US-02370739
PR	02-OCT-2000;	2000US-02370740
PR	13-OCT-2000;	2000US-02399935
PR	13-OCT-2000;	2000US-02399937
PR	20-OCT-2000;	2000US-02409600
PR	20-OCT-2000;	2000US-02411221
PR	20-OCT-2000;	2000US-0241785
PR	20-OCT-2000;	2000US-0241786
PR	20-OCT-2000;	2000US-02463700
PR	08-NOV-2000;	2000US-02464745
PR	08-NOV-2000;	2000US-02464756
PR	08-NOV-2000;	2000US-02464776
PR	08-NOV-2000;	2000US-02464778
PR	08-NOV-2000;	2000US-0246523
PR	08-NOV-2000;	2000US-0246524
PR	08-NOV-2000;	2000US-0246525
PR	08-NOV-2000;	2000US-0246526
PR	08-NOV-2000;	2000US-0246527
PR	08-NOV-2000;	2000US-0246532
PR	08-NOV-2000;	2000US-0246538







XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35277.  
DE Human; Immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX Homo sapiens.  
OS  
XX WO200157182-A2.  
PN  
XX  
PD  
XX  
PF 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01354.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.

PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0234984.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 08-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.

PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Barash SC, Ruben SM;  
PI WPI; 2001-483426/52.  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
PT metastasis -  
XX  
XX Disclosure; SEQ ID NO 35277; 3071pp + Sequence Listing; English.  
XX  
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patients own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/hematopoietic-related diseases, especially  
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/hematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention.  
XX  
XX Sequence 1341 BP; 419 A; 267 C; 301 G; 354 T; 0 other;

Query Match 6.2%; Score 46.2; DB 22; Length 1341;  
Best Local Similarity 67.2%; Pred. No. 0.019;  
Matches 80; Conservative 0; Mismatches 38; Indels 1; Gaps 1;  
QY 617 TTATTAAATTAATTAACATGGTACAAATTTGCAATATAGTGAAGTGTGCAAAATCTCGCTA 676  
Db 1328 TTGTAGAACAATTTATTGGAAGTATATATTTACATA-AGTCAAAATGCATAAATCTTAAGTT 1270  
QY 677 CATGCTCAATGAGTTTTCATATATATTTCCACCCATGTAATCACCACCGAGATCTACA 735  
Db 1269 AAATTTACATGAATTTTACGTAATATATACATCATGTAAGTCACCCACCGAGTCAAGA 1211

RESULT 8  
ABL70463/c  
ID ABL70463 standard; DNA; 6012 BP.  
XX  
XX ABL70463;  
XX  
XX 01-JUL-2002 (first entry)  
XX  
XX Chemically treated cell signalling DNA sequence#177.  
XX  
XX Cell signalling; cytosine methylation; cell signalling disease;  
KW cancer; tumour; cytostatic; ds.  
XX  
XX Unidentified.  
XX  
XX WO200202807-A2.  
XX  
XX 10-JAN-2002.  
XX  
XX 29-JUN-2001; 2001WO-EP07471.  
XX  
XX 30-JUN-2000; 2000DE-1032529.  
PR 01-SEP-2000; 2000DE-1043826.

XX (EPIG-) EPIGENOMICS AG.  
XX PA  
XX Olek A, Piepenbrock C, Berlin K;  
XX WPI; 2002-154758/20.  
XX  
XX Nucleic acid, useful for diagnosis and therapy of diseases associated  
PT with cell signalling e.g. cancer, comprises chemically modified genomic  
PT sequences of genes associated with cell signalling -  
XX  
XX Claim 1; SEQ ID NO 353; 24pp-sequence listing; English.  
XX  
XX The invention relates to a nucleic acid comprising a sequence of at least  
CC 18 bases of a segment of chemically pretreated DNA of genes associated  
CC with cell signalling. The activity of the modified sequences of the  
CC invention may be described as cytostatic. The object of the invention is  
CC to provide the chemically modified DNA of genes associated with cell  
CC signalling, as well as oligonucleotides and/or PNA-oligomers for  
CC detecting cytosine methylations, as well as a method which is  
CC particularly suitable for the diagnosis and/or therapy of genetic and  
CC epigenetic parameters of genes associated with cell signalling. The  
CC chemically modified DNA provided by the invention is useful for diagnosis  
CC and therapy of diseases such as solid tumours and cancer. The sequences  
CC given in records ABL70111-ABL70626 represent chemically pre-treated  
CC genomic DNA's of genes associated with cell signalling.  
CC Note: The sequence data for this patent is not represented in the printed  
CC specification, but is based on sequence information supplied by the  
CC European Patent Office.  
XX  
XX Sequence 6012 BP; 1572 A; 95 C; 1468 G; 2877 T; 0 other;

Query Match 6.1%; Score 45.8; DB 24; Length 6012;  
Best Local Similarity 63.9%; Pred. No. 0.043;  
Matches 85; Conservative 0; Mismatches 47; Indels 1; Gaps 1;  
QY 588 CTTTTCAGAGATTATTTTAAATTAACCTTATTTAAATAATTAACATGGTACAATTTG 647  
Db 1330 CTCATTAATAAATAATATTTTCAATAAATACTATTTACTTATATATAAATACAA-TTA 1272  
QY 648 CATATAGTGAAGTGTGCAAAATCTTCGCTACATGGCTCAATGAGTTTTCATATATTTCC 707  
Db 1271 CATATAATAAATACACAAATCTTAAATAATACAATTCATTAATTTTACAAATATATAC 1212  
QY 708 ACCCATGTAATCA 720  
Db 1211 ACCCATATAACCA 1199

RESULT 9  
ABK31490/c  
ID ABK31490 standard; DNA; 6012 BP.  
XX  
XX ABK31490;  
XX  
XX 23-APR-2002 (first entry)  
XX  
XX Signal transduction associated gene modified DNA #167.  
XX  
XX Human; signal transduction associated gene; cytosine methylation state;  
KW CpG island; signal transduction associated disease; solid tumour; cancer;  
KW antitumour; cytostatic; mutant; ds.  
XX  
XX Homo sapiens.  
OS Synthetic.  
XX  
XX WO200200926-A2.  
XX  
XX 03-JAN-2002.  
XX  
XX 29-JUN-2001; 2001WO-EP07472.  
XX  
XX 30-JUN-2000; 2000DE-1032529.  
PR

PR 01-SEP-2000; 2000DE-1043826.  
XX (EPIC-) EPIGENOMICS AG.  
XX Olek A, Piepenbrock C, Berlin K;  
XX WPI; 2002-147896/19.  
XX  
PT Oligonucleotide for diagnosis and therapy of diseases associated with  
PT signal transduction e.g. cancer, comprises chemically modified genomic  
PT sequences of genes associated with signal transduction -  
XX  
XX Claim 1; SEQ ID NO 333; 24pp; English.  
XX  
XX The present invention relates to chemically modified DNA sequences of  
XX signal transduction associated genes. The DNA sequences are chemically  
XX modified using a solution of bisulphite, hydrogen sulphite or  
XX disulphite. Also disclosed are oligonucleotides and/or PNA oligomers  
XX for detecting the cytosine methylation state (CpG islands) of these  
XX genes, and a method for the diagnosis and/or therapy of genetic and  
XX epigenetic parameters of genes associated with signal transduction.  
XX The genomic DNA can be obtained from cells or cellular components which  
XX contain DNA, e.g. cell lines, biopsies, blood, sputum, stool, urine,  
XX cerebral-spinal fluid, tissue embedded in paraffin such as tissue from  
XX eyes, intestine, kidney, brain, heart, prostate, lung, breast or liver,  
XX histologic object slides, and all their possible combinations. The  
XX sequences of the invention are useful for the diagnosis and therapy of  
XX diseases associated with signal transduction e.g. solid tumours and  
XX cancer. ABK31158-ABK31545 represent chemically pretreated genomic DNA  
XX sequences of different genes associated with signal transduction, or  
XX their complementary sequences.  
XX Note: The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from the  
XX European Patent Office.  
XX  
XX Sequence 6012 BP; 1572 A; 95 C; 1468 G; 2877 T; 0 other;  
XX  
Query Match 6.1%; Score 45.8; DB 24; Length 6012;  
Best Local Similarity 63.9%; Pred. No. 0.043; Indels 1; Gaps 1;  
Matches 85; Conservative 0; Mismatches 47; Indels 1; Gaps 1;  
QY 588 CTTTTCAGAGATTATTTTAAATTAACCTTATTTTAAATAATTAACATGGTACAAATTG 647  
DB 1330 CTCGAATTAATAAATAATTTTTCATAAATACTATTTTACTTATATATAAATAACA-TTA 1272  
QY 648 CATATAGTGAAGTGGCAATCTTCGCTACATGGCTCAATGAGTTTTCATATATATTTCC 707  
DB 1271 CATATATAAATACACAATCTTTAAATAATCAATTCATTAATTTTAAATAATATAC 1212  
QY 708 ACCCATGTAATCA 720  
DB 1211 ACCCATATAACCA 1199  
RESULT 10  
AAL20997  
ID AAL20997 standard; cDNA; 809 BP.  
XX AAL20997;  
XX  
XX 07-DEC-2001 (first entry)  
XX  
DE Human breast cancer expressed polynucleotide 13454.  
XX  
KW Human; breast cancer; cell marker; cytostatic; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200151628-A2.  
XX  
PD 19-JUL-2001.  
XX  
PF 10-JAN-2001; 2001WO-US00798.

XX 14-JAN-2000; 2000US-0176077.  
PR 14-MAR-2000; 2000US-0189167.  
PR 24-MAR-2000; 2000US-0192099.  
PR 29-MAR-2000; 2000US-0193480.  
PR 15-MAY-2000; 2000US-0205230.  
PR 09-JUN-2000; 2000US-0211315.  
PR 25-JUL-2000; 2000US-0220534.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
XX Lillie J, Xu Y, Wang Y, Steinmann K;  
XX WPI; 2001-451856/48.  
XX  
DR New peptide useful as a marker for the diagnosis of breast cancer -  
XX  
XX Claim 1; Page 2389; 3695pp; English.  
XX  
XX The invention relates to human breast cancer expressed polynucleotides  
XX (AAL07544-AAL26789) and methods of assessing whether a patient is  
XX afflicted with breast cancer by examining the correlation between the  
XX expression of certain markers and the cancerous state of breast cells.  
XX The polynucleotides and encoded polypeptides are potential markers for  
XX detecting, diagnosing, monitoring, characterising treating and  
XX potentially preventing breast cancer. The polynucleotides and encoded  
XX polypeptides are also useful for isolating compounds with cytostatic  
XX activity.  
XX  
SQ Sequence 809 BP; 193 A; 213 C; 188 G; 212 T; 3 other;  
XX  
Query Match 6.1%; Score 45.6; DB 22; Length 809;  
Best Local Similarity 63.9%; Pred. No. 0.023;  
Matches 69; Conservative 0; Mismatches 39; Indels 0; Gaps 0;  
QY 643 ATTTCATATAGTGAAGTGTGCAAAATCTTCGCTACATGGCTCAATGAGTTTTCATATA 702  
DB 388 ATTGAATATACACAAGTGTCAAAATCATCAAGTGCACAGCTCGATGAGTTTTCACGACG 447  
QY 703 TTTCACCCCATGTAAATACACCGAGATCTAGATAGTAAGTATGTTTCGT 750  
DB 448 TATACACCGTGTAGTACCGCCGCCGCAACAAGCAAGAAGTGTTCCTT 495  
RESULT 11  
AAS39742  
ID AAS39742 standard; DNA; 3619 BP.  
XX  
XX AAS39742;  
XX  
XX 17-DEC-2001 (first entry)  
XX  
DE Genomic sequence #161 encoding human colon associated polypeptide.  
XX  
XX Human; colon cancer; congenital abnormality; infection; colitis;  
KW inflammatory bowel disease; IBD; neoplastic disorder; gene therapy;  
KW intestinal inflammatory disorder; malabsorption syndrome; gastric;  
KW sigmoid disease; antibacterial; antiviral; antiinflammatory;  
KW cytostatic; ds.  
XX  
OS Homo sapiens.  
XX  
XX WO200155302-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01240.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 20-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216680.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.

PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
  
(HUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Barash SC, Ruben SM;  
WPI; 2001-465567/50.  
  
Isolated polypeptide for treating, preventing and/ or prognosing  
disorders related to the colon including colon cancers and also for  
testing and detection e.g. diagnosis -  
  
Disclosure; SEQ ID No 639; 562pp; English.  
  
The present invention relates to the isolation of novel human colon  
associated polypeptides (AAU22468-AAU22701), and the cDNA and genomic

CC

CC sequences encoding for them. The sequences of the invention are useful  
CC in the diagnosis, treatment, prevention and/or prognosis of disorders  
CC of the colon including colon cancer, congenital abnormalities  
CC (e.g. atresia and stenosis), bacterial and viral infections,  
CC inflammatory bowel disease (IBD), neoplastic cell disorders,  
CC (e.g. polyps and adenomas), intestinal inflammatory disorders, colitis,  
CC colonic inflammation, diarrhoea and dysentery, malabsorption syndromes  
CC (e.g. lactose intolerance), intestinal obstruction and sigmoid diseases.  
CC The polynucleotides sequences of the invention can also be used in gene  
CC therapy. AAS39582-AAS40060 represent DNA sequences encoding for the  
CC novel human colon associated polypeptides of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 3619 BP; 928 A; 860 C; 872 G; 959 T; 0 other;

Query Match 6.1%; Score 45.4; DB 22; Length 3619;  
Best Local Similarity 63.1%; Pred. No. 0.046;  
Matches 70; Conservative 0; Mismatches 41; Indels 0; Gaps 0;  
QY 617 TTATTTAAATATTAACATGGTACATTTTCATATAGTGTGCAAACTTCGCTA 676  
DB 560 TTGTTAGCATTTTGTAAAGATACAGCTTTCAGACAGTAATACGACAGATCTTCAGTG 619  
QY 677 CATGGCTCAATGAGTCTTTTACATATATTTCCACCCATGTAATCACCACCGA 727  
DB 620 CAAAAGTCAATGATGTGACATATGTGTACCTCCCTGTATCATCACTGA 670

RESULT 12  
AAK90083  
ID AAK90083 standard; DNA; 3619 BP.  
AC AAK90083;  
XX  
DT 05-NOV-2001 (first entry)  
XX  
DE Human digestive system antigen genomic sequence SEQ ID NO: 3659.  
XX  
KW Human; digestive system antigen; gene therapy; cancer; appendicitis;  
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;  
KW digestive system disorder; Meckel's diverticulum; ds.  
OS Homo sapiens.  
XX  
PN WO200155314-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01324.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180528.  
PR 24-FEB-2000; 2000US-0184564.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 07-JUN-2000; 2000US-0209467.  
PR 19-MAY-2000; 2000US-0205515.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.

PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0248525.  
PR 08-NOV-2000; 2000US-0248526.  
PR 08-NOV-2000; 2000US-0248527.  
PR 08-NOV-2000; 2000US-0248528.  
PR 08-NOV-2000; 2000US-0248532.  
PR 08-NOV-2000; 2000US-0248609.  
PR 08-NOV-2000; 2000US-0248610.  
PR 08-NOV-2000; 2000US-0248611.  
PR 08-NOV-2000; 2000US-0248613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-502630/55.  
XX Polynucleotides encoding digestive system antigens, useful for  
PT diagnosing, treating, preventing and/or prognosing disorders of the  
PT digestive system, particularly cancer and cancer metastases -  
XX  
XX Disclosure; SEQ ID NO 3659; 986pp; English.  
XX The present invention provides the protein and coding sequences of a  
CC number of human digestive system antigens. These can be used in the  
CC diagnosis, treatment and prevention of digestive system disorders,  
CC including cancer, Meckel's diverticulum, bacterial or parasitic  
CC infectious, appendicitis, Hirschsprung's disease, chronic colitis or  
CC ulcerative colitis. The present sequence is a genomic DNA fragment  
CC encoding a digestive system antigen of the invention.  
XX  
XX Sequence 3619 BP; 928 A; 860 C; 872 G; 959 T; 0 other;  
Query Match 6.1%; Score 45.4; DB 22; Length 3619;  
Best Local Similarity 63.1%; Pred. No. 0.046;  
Matches 70; Conservative 0; Mismatches 41; Indels 0; Gaps 0;  
QY 617 TTATTTAAATATTAACATGTCACAAATTGGCATATAGTGAAGTGCACAAATCTTCCTA 676  
DB 560 TTCTTAGCATTTTGAATAAGATACAGCTTTCAGACAGTAATACGACAGATCTTCAGTG 619

QY 677 CATGGCTCAATGAGTGTTCATATATTTCCACCACCATGTAATCACCACGA 727  
DB 620 CAAAAGCAATGATGTGACATATGTACCCCTGTGTAATCATCACTGA 670  
RESULT 13  
AA23520  
ID AAX23520 standard; DNA: 45546 BP.  
XX AAX23520;  
XX  
XX 23-JUN-1999 (first entry)  
DE Human kidney aminopeptidase P genomic DNA fragment 4.  
KW Aminopeptidase; human; Amp: gene therapy; treatment; Amp-deficiency;  
KW prenatal diagnosis; angioedema; antihypertensive agent; atherosclerosis;  
KW arterial stenosis; industrial protein feed; malabsorption syndrome;  
KW proteinaceous waste degradation; additive; immunohistochemistry; ss.  
OS Homo sapiens.  
XX  
XX WO9911799-A2.  
XX  
XX 11-MAR-1999.  
XX  
XX 02-SEP-1998; 98WO-US18426.  
XX  
XX 02-SEP-1997; 97US-0057854.  
XX  
XX (MEDI-) MEDICAL COLLEGE GEORGIA RES INST.  
XX  
XX PI Ryan JW, Sprinkle TJC, Venema RC;  
XX WPI; 1999-205193/17.  
XX Nucleic acid encoding human aminopeptidase P  
PT  
PS Claim 13; Page 165-192; 201pp; English.  
XX  
XX This invention describes the isolation of a novel human aminopeptidase P  
CC (Amp). This protein is used to produce recombinant Amp and can be used  
CC for gene therapy for treating Amp-deficiency conditions. Its fragments  
CC are used as primers and probes to identify patients with homozygous and  
CC heterozygous Amp deficiency, including prenatal diagnosis of patients  
CC defective in Amp are at risk of developing angioedema if treated with  
CC angiotensin-converting enzyme inhibitors), also as antisense inhibitors  
CC in cases of excessive Amp expression. The product of the invention is  
CC also used to identify Amp-expressing sequences in other animals and to  
CC generate transgenic animals, and comparisons of genomic sequences are  
CC used to detect mutations. Amp inhibitors are potentially useful as  
CC antihypertensive agents and to prevent or treat arterial (re)stenosis  
CC or atherosclerosis. The structure of Amp is used to design synthetic  
CC substrates, e.g. for use in Amp assays. Amp, which hydrolyzes N-terminal  
CC imido bonds, can be used to degrade industrial protein feeds to free  
CC amino acids, to degrade proteinaceous wastes, as additives in enzyme  
CC formulations used to treat malabsorption syndrome and for studying its  
CC biological role. Antibodies against Amp are used in immunohistochemical  
XX methods to study Amp distribution.  
XX  
XX Sequence 45546 BP; 12027 A; 11359 C; 11380 G; 10780 T; 0 other;  
Query Match 6.0%; Score 45.2; DB 20; Length 45546;  
Best Local Similarity 58.0%; Pred. No. 0.14;  
Matches 80; Conservative 0; Mismatches 58; Indels 0; Gaps 0;  
QY 598 GATTATTTTAACTTAACTTATTTAAATATTTAAACATGTCATATTCATATAGTGA 557  
DB 39375 GACCTTTTAAATAAAATTTTAAACAGCTTTATTGAGTATAATGAGCTACAATAA 39434  
QY 658 AGTGTCAATCTCGCTACATGCTCAATGAGTTTTCATATATTTCCACCCCATGTAA 717  
DB 39435 ACTGCACATTTTGAAGTATACATATTCATGATGTTTTCATGATGTTTCCCTCCCA 39494



QY 718 TCACCACCGAGATCTAGA 735  
DB 39495 TCATCACACAGTCAAGA 39512

RESULT 14

ABA62819  
ID ABA62819 standard; DNA; 600 BP.  
XX AC ABA62819;  
XX DT 01-FEB-2002 (first entry)  
XX DE Human foetal liver single exon nucleic acid probe #11124.  
XX KW Human: foetal liver; gene expression; single exon nucleic acid probe; ss.  
XX OS Homo sapiens.  
XX PN WO200157277-A2.  
XX PD 09-AUG-2001.  
XX PF 30-JAN-2001; 2001WO-US00669.  
XX PR 04-FEB-2000; 2000US-0180312.  
XX PR 26-MAY-2000; 2000US-0207456.  
XX PR 30-JUN-2000; 2000US-0608408.  
XX PR 03-AUG-2000; 2000US-0632366.  
XX PR 21-SEP-2000; 2000US-0632366.  
XX PR 27-SEP-2000; 2000US-0234687.  
XX PR 04-OCT-2000; 2000US-0236359.  
XX PR 04-OCT-2000; 2000GB-0024263.  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-483447/52.  
XX DR Human genome-derived single exon nucleic acid probes useful for  
XX PT analyzing gene expression in human fetal liver -  
XX PS Claim 1; SEQ ID NO 11124; 639pp + sequence listing; English.  
XX CC The invention relates to a single exon nucleic acid probe for  
XX CC measuring human gene expression in a sample derived from human foetal  
XX CC liver. The single exon nucleic acid probes may be used for predicting,  
XX CC measuring and displaying gene expression in samples derived from human  
XX CC foetal liver. The present sequence is a single exon nucleic acid  
XX CC probe of the invention.  
XX CC Note: The sequence data for this patent did not form part of the  
XX CC printed specification, but was obtained in electronic format directly  
XX CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 600 BP; 169 A; 98 C; 88 G; 245 T; 0 other;

Query Match 5.9%; Score 44; DB 22; Length 600;  
Best Local Similarity 58.3%; Pred. No. 0.055;  
Matches 77; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 604 TTTTAAATTAACCTATTATTAATAATTAACATGGTACAAATTCGATATAGTGAAGTGTG 663  
DB 75 TTTTAAATTAATTTCTTTTTCCTTAAGTGAGGTATAATACACAGATAGGAAAGTATA 134  
QY 664 CAAATCTTCGCTACATGGCTCAATGAGTTTACATATATTTCCACCCATGTAATCAACCA 723  
DB 135 TTGATCTTAAGTATACACCAATGAATTTTACATATGTTGATACACTAGTATAACTACCA 194  
QY 724 CCGAGATCTAGA 735  
DB 195 CTTGACCAAGA 206

RESULT 15

ABA30104  
ID ABA30104 standard; DNA; 600 BP.  
XX AC ABA30104;  
XX DT 23-JAN-2002 (first entry)  
XX DE Probe #8570 for gene expression analysis in human heart cell sample.  
XX KW Human; gene expression; heart; microarray; vascular system; probe;  
XX KW cardiovascular disease; hypertension; cardiac arrhythmia;  
XX KW congenital heart disease; ss.  
XX OS Homo sapiens.  
XX PN WO200157274-A2.  
XX PD 09-AUG-2001.  
XX PF 30-JAN-2001; 2001WO-US00666.  
XX PR 04-FEB-2000; 2000US-0180312.  
XX PR 26-MAY-2000; 2000US-0207456.  
XX PR 30-JUN-2000; 2000US-0608408.  
XX PR 03-AUG-2000; 2000US-0632366.  
XX PR 21-SEP-2000; 2000US-0234687.  
XX PR 27-SEP-2000; 2000US-0236359.  
XX PR 04-OCT-2000; 2000GB-0024263.  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488899/53.  
XX DR Single exon nucleic acid probes for analyzing gene expression in human  
XX PT hearts -  
XX PS Claim 1; SEQ ID No 8570; 530pp; English.  
XX CC The present invention relates to single exon nucleic acid probes for  
XX CC measuring human gene expression in a sample derived from human heart. The  
XX CC present sequence is one such probe. The probes may be used for  
XX CC predicting, measuring and displaying gene expression in samples derived  
XX CC from the human heart via microarrays. By measuring gene expression, the  
XX CC probes are useful for predicting, diagnosing, grading, staging,  
XX CC monitoring and prognosing diseases of the human heart and vascular system  
XX CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and  
XX CC congenital heart disease.  
XX CC Note: The sequence data for this patent did not form part of the printed  
XX CC specification, but was obtained in electronic format directly from WIPO  
XX CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 600 BP; 169 A; 98 C; 88 G; 245 T; 0 other;

Query Match 5.9%; Score 44; DB 22; Length 600;  
Best Local Similarity 58.3%; Pred. No. 0.055;  
Matches 77; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 604 TTTTAAATTAACCTATTATTAATAATTAACATGGTACAAATTCGATATAGTGAAGTGTG 663  
DB 75 TTTTAAATTAATTTCTTTTTCCTTAAGTGAGGTATAATACACAGATAGGAAAGTATA 134  
QY 664 CAAATCTTCGCTACATGGCTCAATGAGTTTACATATATTTCCACCCATGTAATCAACCA 723  
DB 135 TTGATCTTAAGTATACACCAATGAATTTTACATATGTTGATACACTAGTATAACTACCA 194  
QY 724 CCGAGATCTAGA 735  
DB 195 CTTGACCAAGA 206

Search completed: April 11, 2003, 20:09:56  
Job time : 257.63 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 15:35:54 ; Search time 834.967 Seconds  
(without alignments)  
14547.424 Million cell updates/sec

Title: US-09-914-152-3\_COPY\_1\_750

Perfect score: 750

Sequence: 1 cgcctctgcaaggtagacc.....ctagaatagaatgtattcgt 750

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estnu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_htc:\*
- 9: gb\_estli:\*
- 10: gb\_est2:\*
- 11: gb\_htc:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: gb\_gss:\*
- 18: em\_gss\_hum:\*
- 19: em\_gss\_inv:\*
- 20: em\_gss\_pln:\*
- 21: em\_gss\_vrt:\*
- 22: em\_gss\_fun:\*
- 23: em\_gss\_mam:\*
- 24: em\_gss\_mus:\*
- 25: em\_gss\_other:\*
- 26: em\_gss\_pro:\*
- 27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	7.2	394	9	AI002519
2	53.4	7.1	495	9	AA812008
3	53.4	7.1	532	9	AA761040
4	53.4	7.1	561	9	AA732701
5	52.2	7.0	416	12	BG203593
6	51.8	6.9	502	9	AA768411

7	51.6	6.9	683	17	AG145228
8	51.2	6.8	414	14	BM741058
9	51	6.8	437	9	AI028500
10	50.6	6.7	398	17	AO211285
11	50.4	6.7	576	14	BQ183147
12	49.8	6.6	524	17	CNS01090
13	49.2	6.6	678	17	AG036150
14	47.4	6.3	375	17	B68708
15	46.8	6.2	655	17	BH497623
16	46.6	6.2	257	17	AQ472866
17	46.6	6.2	420	14	H88150
18	46.4	6.2	177	9	AI633412
19	46.2	6.2	245	9	AI302295
20	46.2	6.2	296	9	AI340208
21	46.2	6.2	426	10	BE502085
22	46.2	6.2	474	10	AW055095
23	46.2	6.2	704	17	AG165240
24	45.8	6.1	398	17	AQ231876
25	45.8	6.1	445	17	AQ580954
26	45.8	6.1	1101	17	CNS0039G
27	45.6	6.1	402	13	BM160058
28	45.6	6.1	626	10	AV731485
29	45.6	6.1	743	17	AQ384811
30	45.2	6.0	395	17	AQ526993
31	45.2	6.0	899	17	AQ739502
32	45.2	6.0	911	17	AQ739143
33	44.8	6.0	792	9	AL041300
34	44.6	5.9	571	17	BH720088
35	44.6	5.9	232	9	AI335220
36	44.4	5.9	472	17	B43810
37	44.4	5.9	501	17	AQ234589
38	44.4	5.9	536	17	AQ547304
39	44.2	5.9	239	14	BQ328782
40	44.2	5.9	615	17	AG106461
41	44	5.9	455	17	AQ812510
42	44	5.9	684	17	AG185807
43	43.8	5.8	445	10	AV594211
44	43.6	5.8	419	14	N25194
45	43.6	5.8	539	9	AA702463

ALIGNMENTS

RESULT 1  
AI002519  
LOCUS  
DEFINITION  
394 bp mRNA linear EST 10-NOV-1998  
Oq96h03.s1 NCI-CGAP\_Kid6 Homo sapiens cDNA clone IMAGE:1594229 3'  
similar to contains Alu repetitive element; contains MER36.t3 MER36  
repetitive element ;, mRNA sequence.  
ACCESSION  
AI002519  
VERSION  
AI002519.1  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 394)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
AUTHORS  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL  
Unpublished (1997)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-r@mail.nih.gov](mailto:cgaps-r@mail.nih.gov)  
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: Stratagene, Inc.  
DNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Insert length: 720 Std Error: 0.00

Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 371.  
Location/Qualifiers  
1. .394

FEATURES  
source

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1594229"  
/clone\_lib="NCI\_CGAP\_K1d6"  
/sex="mixed"  
/tissue\_type="kidney tumor"  
/lab\_host="SOLR (kanamycin resistant)"  
/note="Organ: kidney; Vector: Bluescript SK-; Site\_1:  
EcoRI; Site\_2: XhoI; Cloned unidirectionally. Primer:  
Oligo dt. Pooled kidney tumors. 5' adaptor sequence: 5'  
GAATCGGCAGAG 3' 3' adaptor sequence: 5'  
CTCGAGTTTTTTTTTTTTTTT 3' Average insert size: 1.0 kb."  
BASE COUNT 110 a 107 c 72 g 105 t  
ORIGIN

Query Match 7.2%; Score 54; DB 9; Length 394;  
Best Local Similarity 68.2%; Pred. No. 0.0016;  
Matches 75; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
Qy 626 ATATTACATGGTACAAATTCATATAGTGAAGTGTGCAATCTTCGCTACATGGCTCA 685  
Db 272 AACTTTGTTAAAGTACAAATTTATATACATAAATGTAACCAATTTAAATATATAGTTCA 331  
Qy 686 ATGAGTTTTACATATATCCACCATGTAATCACCACGAGATCTAGA 735  
Db 332 ATGGGTTTAAACAATCTATACACCTTGTAAACCAACCACCAATCAAGA 381

RESULT 2  
AA812008  
LOCUS  
DEFINITION  
OB39e12.s1 NCI\_CGAP\_GCB1 Homo sapiens cDNA clone IMAGE:1333774 3'  
Similar to gb:L09753 CD30 LIGAND (HUMAN);, mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 495)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL  
COMMENT  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,  
Ph.D., Gerald Marti, M.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html  
Insert Length: 1467 Std Error: 0.00  
Seq primer: -40ml3 fwd. ET from Amersham.  
High quality sequence stop: 371.  
Location/Qualifiers  
1. .495

FEATURES  
source  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1333774"  
/clone\_lib="NCI\_CGAP\_GCB1"  
/tissue\_type="germinal center B cell"  
/lab\_host="DH10B"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was prepared from human tonsillar cells enriched for

germinal center B cells by flow sorting (CD20+, IgD-),  
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman  
(NCI) and Dr. Gerald Marti (CBER). CDNA synthesis was  
primed with a Not I - oligo(dT) primer  
[5'-TGTTACCAATCTGAAGTGGAGCGGCCCTCATTTTTTTTTTTT-3'  
]. Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT73 vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo."  
BASE COUNT 138 a 102 c 77 g 178 t  
ORIGIN

Query Match 7.1%; Score 53.4; DB 9; Length 495;  
Best Local Similarity 59.6%; Pred. No. 0.0024;  
Matches 90; Conservative 0; Mismatches 61; Indels 0; Gaps 0;  
Qy 600 TTTATTTTAAATTAACCTTTTAAATAATTAACATGCTACAAATTTGCATATAGTGAAG 659  
Db 4 TTTTTTTTTTAAATTTAAATAATTAATTTTAAAGACATGATTTACATACACAAAA 63  
Qy 660 TGTGCAATCTTCGCTACATGGTCAATGAGTTTTTACATATATTTCCACCCATGTAATC 719  
Db 64 TGTATACATTTTAAAGTGTACAGTTTCAGTGAATTTTGACTAATGTACACACCTGTGTCATC 123  
Qy 720 ACCACCGAGATCTAGATAGATGATTCGT 750  
Db 124 ACCACCCCAATGATATTAGAACATTTTCAT 154

RESULT 3  
AA761040  
LOCUS  
DEFINITION  
ny13b04.s1 NCI\_CGAP\_GCB1 Homo sapiens cDNA clone IMAGE:1271599 3'  
Similar to gb:L09753 CD30 LIGAND (HUMAN);, mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 532)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL  
COMMENT  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,  
Ph.D., Gerald Marti, M.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html  
Insert Length: 1408 Std Error: 0.00  
Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 481.  
Location/Qualifiers  
1. .532  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1271599"  
/clone\_lib="NCI\_CGAP\_GCB1"  
/tissue\_type="germinal center B cell"  
/lab\_host="DH10B"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was prepared from human tonsillar cells enriched for  
germinal center B cells by flow sorting (CD20+, IgD-),



QY	600	TTTATTTTTAAATAAACCTTATTTAAAAATATTAAACATGGGTACAAATTTGCATATAGTGAAG	659
Db	4	TTTTGTTTTTTTAAATTTTTAAAAATAAAATTTTATTAAAGACATGATTTTACATACACAAAA	63
QY	660	TGTGCAATCTCGCTACATGGCTCAATGAGTTTTTACATATATTTTCACCATGTAATC	719
Db	64	TGTATACATTTTAAAGTGCAGTTTCAGTGAATTTTGACTAACTACACACCTGTGTGCATC	123
QY	720	ACCACGAGATCTAGAAATAGAATGTATTCGT	750
Db	124	ACCACCCCAATGATATTTAGACATTTTCAT	154

RESULT 7	
AG145228	
LOCUS	683 bp DNA linear GSS 08-JAN-2002
DEFINITION	Pan troglodytes DNA, clone: RP43-006H10.TJ, genomic survey sequence.
ACCESSION	AG145228
VERSION	AG145228.1
KEYWORDS	GI:16674906
SOURCE	GSS.
	Pan troglodytes male lymphocytes DNA, clone_lib:RPCi-43 Chimpanzee Male BAC Library clone:RP43-006H10.TJ.

AUTHORS	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Toki, Y., Watanabe, H. and Sakaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimpesgsc@riken.go.jp, URL: <a href="http://hgpc.gsc.riken.go.jp/">http://hgpc.gsc.riken.go.jp/</a> , Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT	Clones are derived from the chimpanzee BAC library RPCI-43. This BAC end was generated during the Rad process and may have higher chance of clone tracking errors.

```

Sequencing: TJ
LIBRARY
Vector      : pBACE3.6
R.Site 1   : ECORI
R.Site 2   : ECORI
Location/Qualifiers
1..683
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="RP43-006H10.TJ"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RCCI-43 Chimpanzee Male BAC Library"
194 a      128 c      143 g      218 t
BASE COUNT
ORIGIN

```

666 AATCTTCGCTACATGCGTCAATGAGTTTTACATATATTCCACCCATGTAATCACCACC 725  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000

QY 726 GAGATCTAGA 735  
| | | | |  
Db 458 CAGTTCAGA 467

RESULT 8  
BM741058  
LOCUS K-EST0013427 S6SNU620 Homo sapiens cDNA clone S6SNU620-2-G07 5',  
mRNA sequence.  
BM741058  
ACCESSION BM741058.1 GI:19062387  
VERSION  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 414)  
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.  
21C Frontier Korean EST Project 2001  
Unpublished (2002)  
Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 2 row: G column: 07  
High quality sequence stop: 414.  
Location/Qualifiers  
1..414  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="S6SNU620-2-G07"  
/clone\_lib="S6SNU620"  
/sex="F"  
/tissue\_type="Ascites"  
/cell\_type="Scattering floating"  
/cell\_line="SNU-620"  
/lab\_host="Top10F"  
/note="Organ: Stomach; Vector: pcNS; Site\_1: EcoRI;  
Site\_2: NotI; The poly (A)+ RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then decapped  
with tabacco acid pyrophosphatase (TAP). The decapped  
intact mRNA was ligated with DNA-RNA linker including EcoR  
I site by treatment of T4 RNA ligase and the first strand  
cDNA was synthesized from oligo dr-selected mRNA by  
priming with dr-tailed vector. The dr-tailed vector was  
adjusted to have about 60nt. The cDNA vector was  
circularized with E. coli DNA ligase after digestion of  
EcoRI which site is also included in vector. An RNA strand  
converted to a DNA strand by Okayama-Berg method. The  
obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10F, by electroporation method.  
The cDNA libraries constructed by this method are  
full-length enriched cDNA library."

BASE COUNT 136 a 71 c 78 g 129 t  
ORIGIN

Query Match 6.8%; Score 51.2; DB 14; Length 414;  
Best Local Similarity 61.0%; Pred. No. 0.0087;  
Matches 83; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 608 TAATTAACATTTTAAATAATTAAACATGGTACAAATTTGCATATAGTGAAGTGTGCAA 667  
| | | | |  
Db 12 TAATTAATTTCTTTTAAAGTTTCCATGAGTGTGATATACATACAGCAAGTGCACAG 71  
| | | | |

QY 668 TCTTCGTACATGCTCAATCAGTTTTTACATATATTTCCACCCCATGTATCACCACCGA 727  
| | | | |  
Db 72 TCTTAGATATACACTCAGTGAAGCTTCAGAAATGCATGCGATGTAGCCATGATCTCA 131  
| | | | |

QY 728 GATCTAGAATAGAATG 743  
| | | | |  
Db 132 GATCAGATACGCGAG 147  
| | | | |

RESULT 9  
AI028500/c  
LOCUS AI028500.1  
DEFINITION OW44a01.xl Soares\_parathyroid\_tumor\_NbHPA Homo sapiens cDNA clone  
IMAGE:1649640 3' similar to contains Alu repetitive element; , mRNA  
sequence.  
AI028500  
ACCESSION AI028500.1 GI:3245809  
VERSION  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 437)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo  
Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/dbbrp/image/image.html  
Insert Length: 1520 Std Error: 0.00  
Seq primer: -40ml3 fwd. Et from Amersham  
High quality sequence stop: 417.  
Location/Qualifiers  
1..437  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1649640"  
/clone\_lib="Soares\_parathyroid\_tumor\_NbHPA"  
/tissue\_type="parathyroid tumor"  
/dev\_stage="adult"  
/lab\_host="DH10B (ampicillin resistant)"  
/note="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia  
) with a modified polylinker; Site\_1: Not I; Site\_2: Eco  
RI; 1st strand cDNA was primed with a Not I - oligo(dT)  
primer  
15'-TGTTCACCAATCTGAAGTGGGAGCGCGCCGACCAATTTTTTTTTTTTTTTT  
TTTTT-3', double-stranded cDNA was size selected, ligated  
to Eco RI adapters (Pharmacia), digested with Not I and  
cloned into the Not I and Eco RI sites of a modified pT7T3  
vector (Pharmacia). Library went through one round of  
normalization to a Cot = 5. Library constructed by Bento  
Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid  
adenomas was kindly provided by Dr. Stephen Marx, National  
Institute of Diabetes and Digestive and Kidney Diseases,  
NIH."

BASE COUNT 104 a 106 c 86 g 141 t  
ORIGIN

Query Match 6.8%; Score 51; DB 9; Length 437;  
Best Local Similarity 61.8%; Pred. No. 0.0099;  
Matches 81; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 606 TTTAATTAACATTTTAAATAATTAAACATGGTACAAATTTGCATATAGTGAAGTGTGCA 665  
| | | | |  
Db 423 TATCATTTAAATAATTATATAACTTATTAAAGTATGATCTCCATACAGACAGTATACA 364  
| | | | |

QY 666 AATCTCGGTACATGCTCAATCAGTTTTTACATATATTTCCACCCCATGTATCACCACC 725  
| | | | |  
Db 363 AATTTTAAGTATACAGCTTGATGAATTTATACAAAGTGAACATATCATCTTAACCTCCAGC 304  
| | | | |

[illegible]



[illegible][illegible]

```

Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

FEATURES             Location/Qualifiers
     source            1..655
                        /organism="Brassica oleracea"
                        /strain="TO1000DH3"
                        /db_xref="taxon:3712"
                        /clone="BOGOQ27"
                        /clone_lib="BOGO"
                        /note="Vector: pHS1; site_1: BstXI; 2-3 kb sheared
                        genomic DNA inserted into pHS1 using BstXI linkers"
BASE COUNT          242 a      87 c      90 g      236 t
ORIGIN
Query Match          6.2%; Score 46.8; DB 17; Length 655;
Best Local Similarity 58.7%; Pred. No. 0.13;
Matches 81; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

```

	maches	oi,	conservative	o,	flamenco	377,	indels	o,	gaps
Qy	604	TTTTTAATTAAACTTTATTTAAAAAATATTACATGGTACAATTTGCCATATAGTGAAGTG	663						
Db	640	TCITAAAATAAAATTTATTAACCAAAAATATGTGAAATTTATATAAACTATAATATA	581						
Qy	664	CAAATCTTCGCTACATGGGTCATAGTTTTTTACATATATTTCCACCATTGTAATCACCA	723						
Db	580	TAACTTATATATAATTATTTAAATATATGTATATATATATGCATATAAATCATCAAAT	521						
Qy	724	CCGAGATCTAGAAATAGAA	741						
Db	520	TGGATATCCGTACTATAA	503						

Search completed: April 11, 2003, 22:37:39  
Job time : 842.967 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 15:37:18 ; Search time 20.7101 Seconds  
(without alignments)  
11106.080 Million cell updates/sec

Title: US-09-914-152-3\_COPY\_1\_750

Perfect score: 750  
Sequence: 1 cgcctctgcaagtagacc.....ctagaatagaatgatttcgt 750

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45.2	6.0	45546	4	US-09-146-053-6
2	42.6	5.7	50000	4	US-09-146-053-3
3	37.6	5.0	11299	4	US-09-238-356-14
4	37.6	5.0	36651	4	US-09-738-894A-3
5	37.4	5.0	9171	1	US-08-038-662-5
6	37.4	5.0	9171	1	US-08-302-832-5
7	37.4	5.0	9171	2	US-08-530-198-5
8	37.4	5.0	9171	2	US-08-469-880-5
9	37.4	5.0	9171	2	US-08-728-470-5
10	37.4	5.0	9171	2	US-08-617-697-5
11	37.4	5.0	9171	4	US-08-719-641-5
12	37.4	5.0	9323	1	US-08-038-682-6
13	37.4	5.0	9323	1	US-08-302-832-6
14	37.4	5.0	9323	2	US-08-530-198-6
15	37.4	5.0	9323	2	US-08-469-880-6
16	37.4	5.0	9323	2	US-08-728-470-6
17	37.4	5.0	9323	2	US-08-617-697-6
18	37.4	5.0	9323	4	US-08-719-641-6
19	37.2	5.0	98844	4	US-09-791-211-10
20	36.2	4.8	50000	4	US-09-146-053-4
21	35.8	4.8	1750	4	US-09-276-531-34
22	35.8	4.8	162450	4	US-09-345-882-1
23	35.4	4.7	6623	2	US-08-687-080-68
24	35.2	4.7	2368	3	US-08-714-918-2
25	35.2	4.7	2368	4	US-09-265-315-2
26	35.2	4.7	2368	4	US-09-265-315-2
27	35.2	4.7	2368	4	US-09-266-417-2

28	35.2	4.7	5526	3	US-08-751-359-21	Sequence 21, Appl
29	35.2	4.7	5526	4	US-08-907-146-21	Sequence 21, Appl
C 30	34.8	4.6	722	4	US-08-998-416-780	Sequence 780, App
C 31	34.8	4.6	727	4	US-08-998-416-1011	Sequence 1011, Ap
C 32	34.6	4.6	2417	4	US-09-629-645A-3	Sequence 3, Appli
C 33	34	4.5	603	4	US-09-134-001C-1226	Sequence 1226, Ap
C 34	34	4.5	4287	1	US-08-038-682-7	Sequence 7, Appli
C 35	34	4.5	4287	1	US-08-302-832-7	Sequence 7, Appli
C 36	34	4.5	4287	2	US-08-530-198-7	Sequence 7, Appli
C 37	34	4.5	4287	2	US-08-469-880-7	Sequence 7, Appli
C 38	34	4.5	4287	2	US-08-728-470-7	Sequence 7, Appli
C 39	34	4.5	4287	4	US-08-719-641-7	Sequence 7, Appli
C 40	33.6	4.5	1215	2	US-09-092-770-8	Sequence 8, Appli
C 41	33.6	4.5	1215	4	US-09-222-851-8	Sequence 8, Appli
C 42	33.2	4.4	2547	3	US-08-262-220-7	Sequence 7, Appli
C 43	33.2	4.4	2547	3	US-08-471-733-7	Sequence 7, Appli
C 44	33.2	4.4	2547	3	US-08-468-878-7	Sequence 7, Appli
C 45	33.2	4.4	2547	4	US-08-750-494-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1  
US-09-146-053-6  
; Sequence 6, Application US/09146053A  
; Patent No.: 6399349  
; GENERAL INFORMATION:  
; APPLICANT: Ryan, James W.  
; APPLICANT: Sprinkle, Terry Joe Curtis  
; APPLICANT: Venema, Richard C.  
; TITLE OF INVENTION: Human Amino peptidase P Gene  
; FILE REFERENCE: MCG103  
; CURRENT APPLICATION NUMBER: US/09/146,053A  
; CURRENT FILING DATE: 1998-09-02  
; EARLIER APPLICATION NUMBER: 60/057,854  
; EARLIER FILING DATE: 1997-09-02  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 45546  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-146-053-6

Query Match	6.0%	Score 45.2;	DB 4;	Length 45546;
Best Local Similarity	58.0%	Pred. No. 0.0039;		
Matches	80;	Conservative	0;	Mismatches 58; Indels 0; Gaps 0;
Qy	598	GATTATTTTAACTTATTTAAATAATTTAAACATGGTACAAATTTGCATATAGTGA	657	
Db	39375	GACCTTTTAAATAATTTAAACAGCTTTATTGAGGTATATGGAGCTACAATAA	39434	
Qy	658	AGTGTGCAATCTCGCTACATGGCTCAATGAGTATTTTACATATATTTCCACCCATGAA	717	
Db	39435	ACTGCACATTTTGAAGTATACAAATTTGACGTTTGGATGTGTTATTTCCCATGAAA	39494	
Qy	718	TCACACGAGATCTAGA	735	
Db	39495	TCATCACCACAGTCAAGA	39512	

RESULT 2  
US-09-146-053-3  
; Sequence 3, Application US/09146053A  
; Patent No.: 6399349  
; GENERAL INFORMATION:  
; APPLICANT: Ryan, James W.  
; APPLICANT: Sprinkle, Terry Joe Curtis  
; APPLICANT: Venema, Richard C.  
; TITLE OF INVENTION: Human Amino peptidase P Gene  
; FILE REFERENCE: MCG103  
; CURRENT APPLICATION NUMBER: US/09/146,053A

```

Db 6713 TTATTTATAGACGAGTTATATGGTTTAAAGCTTCTAATAATAGTAGCAGTATTTATATCA 6654

QY 656 GAAGTGTGCAAACTTTCGCTACATGGCTCAATGAGTGTTCATATATTTCCACCCATGT 715
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6653 TATGCATGCTTCTATTTGGTCTCATACTGAAGTAATTTTGAATGGTCTACACCCGGGA 6594
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 716 AATCACACCGAGATCTAGA 735
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6593 AGGAACCCGAGCATTTCTATA 6574
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 4
US-09-738-894A-3/c
; Sequence 3, Application US/09738894A
; Patent No. 6331423
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO00636
; CURRENT APPLICATION NUMBER: US/09/738,894A
; CURRENT FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 36651
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(36651)
; OTHER INFORMATION: n = A,T,C or G
US-09-738-894A-3

Query Match 5.0%; Score 37.6; DB 4; Length 36651;
Best Local Similarity 56.5%; Pred. No. 0.65;
Matches 70; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 627 ATATTAACATGCTACAAATTTGCATATAGTGAAGTGTGCAAAATCTTCGCTACATGGCTCAA 686
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 24765 ATTTTATGATACATAAATTAATATACAAATACGTTGACAGTTCITAGACATTCAGCTCAA 24706
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 687 TGAGTTTTTACATATATTTCCACCCCATGTAAATCACCCGAGATCTAGAATAGAAATGTAT 746
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 24705 TGAATTTTAAATTAATAATACGCCCTCATAAATACCCCAACCAAGATATAAAACATT 24646
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 747 TCGT 750
      ||
Db 24645 CTGT 24642

RESULT 5
US-08-038-682-5/c
; Sequence 5, Application US/08038682
; Patent No. 5549897
; GENERAL INFORMATION:
; APPLICANT: BARENKAMP, STEPHEN J
; APPLICANT: ST. GEME III, JOSEPH W
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT SURFACE PROTEINS
; TITLE OF INVENTION: OF NON-TYPEABLE HAEMOPHILUS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
; STREET: Bldg. 1
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/038.682
; FILING DATE: 16-MAR-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: BERKSTRESSER, JERRY W
; REGISTRATION NUMBER: 22,651
; REFERENCE/DOCKET NUMBER: 1038-293
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9171 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-038-682-5

Query Match          5.0%; Score 37.4; DB 1; Length 9171;
Best Local Similarity 48.0%; Pred. No. 0.38;
Matches 107; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 498 GTTTGTGAGACACTGCCCATGAATGCTCTCCAGAAATAGTCGGATTGTGGTCAA 557
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 5191 GTTCAAGTGCACGATGAAGCCCTTTTACTAAAGCGCTTCTTCTGCATACAA 5132

QY 558 TAAATTTGGGCAATCTACAGAACATGCTCTTTTCAGAGATTATTTTAAATAACT 617
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 5131 GATGAAGAGCCGCGCAAGCAATGATATGCTGACGTGATTTTAAATCTGCTTTTATA 5072

QY 618 TATTTAAATATTAACATGTTGCAATTTGCAATAGTGAAGTGCACAAATCTTCGCTAC 677
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 5071 TTTTTCATAATAAAGCTGTAAAGTAACTATTATGTTATTCCTCGTAATTTTACAAG 5012

QY 678 ATGGCTCAATGAGTTTACATATATTTCCACCCATGTAATCA 720
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 5011 ATGGGTAAGCCCGTACTGAACTTTAACCCACACAGATAATAA 4969

RESULT 6
US-08-302-832-5/C
; Sequence 5, Application US/08302832
; Patent No. 5603938
; GENERAL INFORMATION:
; APPLICANT: Barenkamp, Stephen J
; TITLE OF INVENTION: High Molecular Weight Surface Proteins
; TITLE OF INVENTION: of No. 5603938-Typeable Haemophilus
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd.
; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
; STREET: Bldg. 1
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/302,832
; FILING DATE: 16-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9205704.1
; FILING DATE: 16-MAR-1992
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US pct/us93/02166
; FILING DATE: 16-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Berkstresser, Jerry W
; REGISTRATION NUMBER: 22,651
; REFERENCE/DOCKET NUMBER: 1038-404
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9171 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-302-832-5

Query Match          5.0%; Score 37.4; DB 1; Length 9171;
Best Local Similarity 48.0%; Pred. No. 0.38;
Matches 107; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 498 GTTTGTGAGACACTGCCCATGAATGCTCTCCAGAAATAGTCGGATTGTGGTCAA 557
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 5191 GTTCAAGTGCACGATGAAGCCCTTTTACTAAAGCGCTTCTTCTGCATACAA 5132

QY 558 TAAATTTGGGCAATCTACAGAACATGCTCTTTTCAGAGATTATTTTAAATAACT 617
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 5131 GATGAAGAGCCGCGCAAGCAATGATATGCTGACGTGATTTTAAATCTGCTTTTATA 5072

QY 618 TATTTAAATATTAACATGTTGCAATTTGCAATAGTGAAGTGCACAAATCTTCGCTAC 677
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 5071 TTTTTCATAATAAAGCTGTAAAGTAACTATTATGTTATTCCTCGTAATTTTACAAG 5012

QY 678 ATGGCTCAATGAGTTTACATATATTTCCACCCATGTAATCA 720
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 5011 ATGGGTAAGCCCGTACTGAACTTTAACCCACACAGATAATAA 4969

RESULT 7
US-08-302-198-5/C
; Sequence 5, Application US/08530198
; Patent No. 5869065
; GENERAL INFORMATION:
; APPLICANT: BARENKAMP, STEPHEN J
; APPLICANT: ST. GEME III, JOSEPH W
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT SURFACE PROTEINS
; TITLE OF INVENTION: OF NON-TYPEABLE HAEMOPHILUS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
; STREET: Bldg. 1
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/530,198
; FILING DATE: 13-DEC-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: BERKSTRESSER, JERRY W
; REGISTRATION NUMBER: 22,651
; REFERENCE/DOCKET NUMBER: JWB-1186
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813

```

```
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9171 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-530-198-5

Query Match          5.0%; Score 37.4; DB 2; Length 9171;
Best Local Similarity 48.0%; Pred. No. 0.38;
Matches 107; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 498 GTTTTGTGACACACTGCCCATGAATGCTCTCCAGAAATAGTCGGATTGTTGGTCAAA 557
Db 5191 GTTTCAGTGCACAGATACTGAAGCCCTTTTACTAAAGCGCTTCTTCGCATACAAT 5132
QY 558 TAAATTTGGGCAATCTACAGAACATGTGCTTTTTCAGAGATTATTTTAAATTAAC 617
Db 5131 GATGAGAACCCAGGCCAAGCAATAGTATTCACAGTGTGTTTAACTGCTTTTATA 5072
QY 618 TATTTAAATAATTAACATGTTGACAAATTTGCATATAGTAGGAGTGCGAAATCTTCGCTAC 677
Db 5071 TTTTTCATATAATAACCTGTGTTAAATAACTTTTATTGTATTCGCGTAATTTTACAAG 5012
QY 678 ATGGCTCAATGAGTTTTCATATATTTCCACCCATGTAATCA 720
Db 5011 ATGGGTAAGCCCGTACTGAACCTTTAAACCCACACAGTAATAA 4969

RESULT 8
US-08-469-880-5/c
; Sequence 5, Application US/08469880
; Patent No. 5876733
; GENERAL INFORMATION:
; APPLICANT: Barenkamp, Stephen J.
; TITLE OF INVENTION: High Molecular Weight Surface Proteins
; TITLE OF INVENTION: of No. 5876733-Typeable Haemophilus
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd.
; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,880
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9205704.1
; FILING DATE: 16-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US93/02166
; FILING DATE: 16-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/302,832
; FILING DATE: 16-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Berkstresser, Jerry W
; REGISTRATION NUMBER: 22,651
; REFERENCE/DOCKET NUMBER: 1038-516
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; INFORMATION FOR SEQ ID NO: 5:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9171 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-469-880-5

Query Match          5.0%; Score 37.4; DB 2; Length 9171;
Best Local Similarity 48.0%; Pred. No. 0.38;
Matches 107; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 498 GTTTTGTGACACACTGCCCATGAATGCTCTCCAGAAATAGTCGGATTGTTGGTCAAA 557
Db 5191 GTTTCAGTGCACAGATACTGAAGCCCTTTTACTAAAGCGCTTCTTCGCATACAAT 5132
QY 558 TAAATTTGGGCAATCTACAGAACATGTGCTTTTTCAGAGATTATTTTAAATTAAC 617
Db 5131 GATGAGAACCCAGGCCAAGCAATAGTATTCACAGTGTGTTTAACTGCTTTTATA 5072
QY 618 TATTTAAATAATTAACATGTTGACAAATTTGCATATAGTAGGAGTGCGAAATCTTCGCTAC 677
Db 5071 TTTTTCATATAATAACCTGTGTTAAATAACTTTTATTGTATTCGCGTAATTTTACAAG 5012
QY 678 ATGGCTCAATGAGTTTTCATATATTTCCACCCATGTAATCA 720
Db 5011 ATGGGTAAGCCCGTACTGAACCTTTAAACCCACACAGTAATAA 4969

RESULT 9
US-08-728-470-5/c
; Sequence 5, Application US/08728470
; Patent No. 5928651
; GENERAL INFORMATION:
; APPLICANT: Barenkamp, Stephen J.
; TITLE OF INVENTION: High Molecular Weight Surface Proteins
; TITLE OF INVENTION: of No. 5928651-Typeable Haemophilus
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd.
; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/728,470
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/302,832
; FILING DATE: 16-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US93/02166
; FILING DATE: 16-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9205704.1
; FILING DATE: 16-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Berkstresser, Jerry W
; REGISTRATION NUMBER: 22,651
; REFERENCE/DOCKET NUMBER: 1038-633
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
```

```
/ LENGTH: 9171 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
US-08-728-470-5

Query Match          5.0%; Score 37.4; DB 2; Length 9171;
Best Local Similarity 48.0%; Pred. No. 0.38;
Matches 107; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 498 GTTTTGTGAGACACTGCCCATGAATGCTCTCCAGAAATAGTCGGATTTGGTCAAA 557
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5191 GTTTCAGAGTCACAGATAACTGAAAGCCTTTTACTAAAAACGCTTCTCTGCATACAAT 5132

QY 558 TAAATTTGGGCAATTCACAGACATGCTCTCTCCAGAAATAGTCGGATTTGGTCAAA 617
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5131 GATGAGAAGCCAGCCGCAACTACTGATATTCGACTGAGTTTAACTGCTTTTATA 5072

QY 618 TATTTAAAAATATTAAACATGTCATATTTGCATATAGTCAAGTGCATATCTTCGCTAC 677
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5071 TTTTTCATATATAACCTGTTAAAAAATCTTTTATTGTTATTCGCTAAATTTTACAAG 5012

QY 678 ATGGCTCAATGAGTTTTCATATATTTCCACCCATGTAATCA 720
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5011 ATGGGTAAGCCCGTACTGAACCTTTAACCCACACAGTAATAA 4969

RESULT 10
US-08-617-697-5/c
; Sequence 5, Application US/08617697
; Patent No. 5977336
; GENERAL INFORMATION:
; APPLICANT: Barenkamp, Stephen J
; TITLE OF INVENTION: High Molecular Weight Surface Proteins
; TITLE OF INVENTION: of No. 5977336-Typeable Haemophilus
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd.
; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
; STREET: Bldg. 1
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/617.697
; FILING DATE: 01-APR-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/302,832
; FILING DATE: 05-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US93/02166
; FILING DATE: 16-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Berkstreser, Jerry W.
; REGISTRATION NUMBER: 22,651
; REFERENCE/DOCKET NUMBER: 1038-557
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9171 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
/ MOLECULE TYPE: DNA (genomic)
US-08-617-697-5

Query Match          5.0%; Score 37.4; DB 2; Length 9171;
Best Local Similarity 48.0%; Pred. No. 0.38;
Matches 107; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 498 GTTTTGTGAGACACTGCCCATGAATGCTCTCCAGAAATAGTCGGATTTGGTCAAA 557
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5191 GTTTCAGAGTCACAGATAACTGAAAGCCTTTTACTAAAAACGCTTCTCTGCATACAAT 5132

QY 558 TAAATTTGGGCAATTCACAGACATGCTCTCTCCAGAAATAGTCGGATTTGGTCAAA 617
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5131 GATGAGAAGCCAGCCGCAACTACTGATATTCGACTGAGTTTAACTGCTTTTATA 5072

QY 618 TATTTAAAAATATTAAACATGTCATATTTGCATATAGTCAAGTGCATATCTTCGCTAC 677
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5071 TTTTTCATATATAACCTGTTAAAAAATCTTTTATTGTTATTCGCTAAATTTTACAAG 5012

QY 678 ATGGCTCAATGAGTTTTCATATATTTCCACCCATGTAATCA 720
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5011 ATGGGTAAGCCCGTACTGAACCTTTAACCCACACAGTAATAA 4969

RESULT 11
US-08-719-641-5/c
; Sequence 5, Application US/08719641
; Patent No. 6218141
; GENERAL INFORMATION:
; APPLICANT: Barenkamp, Stephen J
; TITLE OF INVENTION: High Molecular Weight Surface Proteins
; TITLE OF INVENTION: of No. 6218141-Typeable Haemophilus
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd.
; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
; STREET: Bldg. 1
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/719,641
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/302,832
; FILING DATE: 16-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US93/02166
; FILING DATE: 16-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Berkstreser, Jerry W.
; REGISTRATION NUMBER: 22,651
; REFERENCE/DOCKET NUMBER: 1038-625
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9171 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
```

US-08-719-641-5

Query Match 5.0%; Score 37.4; DB 4; Length 9171;  
Best Local Similarity 48.0%; Pred. No. 0.38;  
Matches 107; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 498 GTTTGTGAGACACTGCCCATGAATGCTCTCCAGAAATAGTCGGATTGTGTCGCAAA 557  
DB 5191 GTTCAAGTGCACAGATACTGAAGCCCTTTTACTAAAAACGGTCTCTCTGCATACAAT 5132

QY 558 TAAATTTGGCAATCTACAGACATGTCTCTTTTTCAGAGATTATTTTAAATTAAC 617  
DB 5131 GATGAAGAGCCAGGCCAAGCAATAGTATTCAGCTGAGTTTAACTGCTTTTATA 5072

QY 618 TATTTAAATAATTAACATGGTACAAATTTGCATATAGTGAAGTGTGCAAAATCTTCGCTAC 677  
DB 5071 TTTTTCATAATAAACCCTGTTAAAAATACCTTTATGTTATTCCTCGTAAATTTTACAAG 5012

QY 678 ATGGCTCAATGAGTTTTCATATATTTCCACCCTCATCA 720  
DB 5011 ATGGGTAAGCCGCTAGTGAACCTTTAACCCACACAGTAATAA 4969

RESULT 12  
US-08-682-6/c  
; Sequence 6, Application US/08038682  
; Patent No. 5549897  
; GENERAL INFORMATION:  
; APPLICANT: BARENKAMP, STEPHEN J  
; APPLICANT: ST. GEME III, JOSEPH W  
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT SURFACE PROTEINS  
; TITLE OF INVENTION: OF NON-TYPEABLE HAEMOPHILUS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Shoemaker and Mattare, Ltd.  
; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza  
; STREET: Bldg. 1  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202-0286

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/038,682  
; FILING DATE: 16-MAR-1993  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BERKSTRESSER, JERRY W  
; REGISTRATION NUMBER: 22,651  
; REFERENCE/DOCKET NUMBER: 1038-293  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 415-0810  
; TELEFAX: (703) 415-0813  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9323 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)

US-08-038-682-6  
Query Match 5.0%; Score 37.4; DB 1; Length 9323;  
Best Local Similarity 48.0%; Pred. No. 0.38;  
Matches 107; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 498 GTTTGTGAGACACTGCCCATGAATGCTCTCCAGAAATAGTCGGATTGTGTCGCAAA 557  
DB 5502 GTTTCAGTGCACAGATACTGAAGCCCTTTTACTAAAAACGGTCTCTCTGCATACAAT 5443

US-08-719-641-5  
Query Match 5.0%; Score 37.4; DB 4; Length 9171;  
Best Local Similarity 48.0%; Pred. No. 0.38;  
Matches 107; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 498 GTTTGTGAGACACTGCCCATGAATGCTCTCCAGAAATAGTCGGATTGTGTCGCAAA 557  
DB 5191 GTTCAAGTGCACAGATACTGAAGCCCTTTTACTAAAAACGGTCTCTCTGCATACAAT 5132

QY 558 TAAATTTGGCAATCTACAGACATGTCTCTTTTTCAGAGATTATTTTAAATTAAC 617  
DB 5131 GATGAAGAGCCAGGCCAAGCAATAGTATTCAGCTGAGTTTAACTGCTTTTATA 5072

QY 618 TATTTAAATAATTAACATGGTACAAATTTGCATATAGTGAAGTGTGCAAAATCTTCGCTAC 677  
DB 5071 TTTTTCATAATAAACCCTGTTAAAAATACCTTTATGTTATTCCTCGTAAATTTTACAAG 5012

QY 678 ATGGCTCAATGAGTTTTCATATATTTCCACCCTCATCA 720  
DB 5011 ATGGGTAAGCCGCTAGTGAACCTTTAACCCACACAGTAATAA 4969

QY 558 TAAATTTGGCAATCTACAGACATGTCTCTTTTTCAGAGATTATTTTAAATTAAC 617  
DB 5442 GATGAAGAGCCAGGCCAAGCAATAGTATTCAGCTGAGTTTAACTGCTTTTATA 5383

QY 618 TATTTAAATAATTAACATGGTACAAATTTGCATATAGTGAAGTGTGCAAAATCTTCGCTAC 677  
DB 5382 TTTTTCATAATAAACCCTGTTAAAAATACCTTTATGTTATTCCTCGTAAATTTTACAAG 5323

QY 678 ATGGCTCAATGAGTTTTCATATATTTCCACCCTCATCA 720  
DB 5322 ATGGGTAAGCCGCTAGTGAACCTTTAACCCACACAGTAATAA 5280

RESULT 13  
US-08-302-832-6/c  
; Sequence 6, Application US/08302832  
; Patent No. 5603938  
; GENERAL INFORMATION:  
; APPLICANT: BARENKAMP, STEPHEN J  
; TITLE OF INVENTION: High Molecular Weight Surface Proteins  
; TITLE OF INVENTION: of No. 5603938-Typeable Haemophilus  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Shoemaker and Mattare, Ltd.  
; STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza  
; STREET: Bldg. 1  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202-0286

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/302,832  
; FILING DATE: 16-SEP-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9205704.1  
; FILING DATE: 16-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US pct/us93/02166  
; FILING DATE: 16-MAR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Berkstresser, Jerry W  
; REGISTRATION NUMBER: 22,651  
; REFERENCE/DOCKET NUMBER: 1038-404  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 415-0810  
; TELEFAX: (703) 415-0813  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9323 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)

US-08-302-832-6  
Query Match 5.0%; Score 37.4; DB 1; Length 9323;  
Best Local Similarity 48.0%; Pred. No. 0.38;  
Matches 107; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 498 GTTTGTGAGACACTGCCCATGAATGCTCTCCAGAAATAGTCGGATTGTGTCGCAAA 557  
DB 5502 GTTTCAGTGCACAGATACTGAAGCCCTTTTACTAAAAACGGTCTCTCTGCATACAAT 5443

QY 558 TAAATTTGGCAATCTACAGACATGTCTCTTTTTCAGAGATTATTTTAAATTAAC 617  
DB 5442 GATGAAGAGCCAGGCCAAGCAATAGTATTCAGCTGAGTTTAACTGCTTTTATA 5383



Qy	618	TATTATAAAATATTAACATAGGTACAAATTTGCAATATAGTGAAGTGTGCAAACTCTTCGTAC	677
Db	5382	TTTTTCATAAATAAACCTGTTAAAAAATACTTTTATGTATCTCCGTAATTTTTTTACAAG	5323
Qy	678	ATGCTCAATGAGTTTTTACATATATTTCCACCCATGTAATCA	720
Db	5322	ATGGTAAAGCCCGTACTGAACTTTAAACCCACACAGTAATAA	5380

## RESULT 7A

US-08-530-198-6/c  
; Sequence 6, Application US/08530198  
; Patent No. 5869065

GENERAL INFORMATION:  
 APPLICANT: BARENKAMP, STEPHEN J  
 APPLICANT: ST. GENE III, JOSEPH W  
 TITLE OF INVENTION: HIGH MOLECULAR WEIGHT SURFACE PROTEINS  
 TITLE OF INVENTION: OF NON-TYPEABLE HAEMOPHILUS  
 NUMBER OF SEQUENCES: 8  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Shoemaker and Mattare, Ltd  
 STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza  
 STREET: Bldg. 1  
 CITY: Arlington  
 STATE: Virginia  
 COUNTRY: U.S.A.  
 ZIP: 22202-0286

```

1  COMPUTER READABLE FORM:
2
3  MEDIUM TYPE: Floppy disk
4
5  COMPUTER: IBM PC compatible
6
7  OPERATING SYSTEM: PC-DOS/MS-DOS
8
9  SOFTWARE: Patent In Release #1.0, Version #1.25
10
11  CURRENT APPLICATION DATA:
12
13  APPLICANT NUMBER: US/08/530,198
14
15  FILING DATE: 13-DEC-1995
16
17  CLASSIFICATION: 424
18
19  ATTORNEY/AGENT INFORMATION:
20
21  NAME: BERKSTRESSER, JERRY W
22
23  REGISTRATION NUMBER: 22,651
24
25  REFERENCE/DOCKET NUMBER: JWB-1186
26
27  TELECOMMUNICATION INFORMATION:
28
29  TELEPHONE: (703) 415-0810
30
31  TELEFAX: (703) 415-0813
32
33  INFORMATION FOR SEQ ID NO: 6:
34
35  SEQUENCE CHARACTERISTICS:
36
37  LENGTH: 9323 base pairs
38
39  TYPE: nucleic acid
40
41  STRANDEDNESS: single
42
43  TOPOLOGY: linear
44
45  MOLECULE TYPE: DNA (genomic)
46
47  US-08-530-198-6

```

Query Match 5.0%; Score 37.4; DB 2; Length 9323;  
Best Local Similarity 48.0%; Pred. No. 0.38;  
Matches 107; Conservative 0; Mismatches 116; Indels 0

498	QY	GTTTGTGAGACACTGGCCCAATGAATGCTCTCCAGAAATAGTCGGATTGTGGTCAAA	557
5502	Db	GTTCRAAGTCACCAGATAACTGAAGCGCTTTTACTAAAGCGCTTCTCTGCATACAAT	5443
558	QY	TAAATTTGGGCAATTCTACAGAACATGTCCTTTTTCAGAGATTATTTTAAATTAACCT	617
5442	Db	GATGAGAAGCCAGGCCCAAGCAATACTGATATATGCACTGAGTTTAAATCTGCTTTTATA	5383
618	QY	TATTTAAAAATATTAACTAGGTACAAATTTGCATATAGTCAAGTGTGCAAAATCTCGCTAC	677
5382	Db	TTTTTCATAATAATAACCTGTTAAATAATCTTATGTATTTCCGTATTTTTCACAAG	5323
678	QY	ATGGCTCAATGAGTTTTTACATATATTTTCCACCAATGTAATCA	720
5322	Db	ATGGGTAAAGCCGCTACTGAACTTTTAAACCCACACAGTAAATAA	5280

RESULT 15

US-08-469-880-6/c  
Sequence 6, Application US/08469880  
Patent No. 5876733  
GENERAL INFORMATION:  
APPLICANT: Barenkamp, Stephen J.  
TITLE OF INVENTION: High Molecular Weight Surface Proteins  
TITLE OF INVENTION: Of No. 5876733-Typeable Haemophilus  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Shoemaker and Mattare, Ltd.  
STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza  
STREET: Bldg. 1  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202-0286  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,880  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9205704.1  
FILING DATE: 16-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US93/02166  
FILING DATE: 16-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/302,832  
FILING DATE: 16-SEP-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Berkstresser, Jerry W  
REGISTRATION NUMBER: 22,651  
REFERENCE/DOCKET NUMBER: 1038-516 MLS:vg  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 415-0810  
TELEFAX: (703) 415-0813  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9323 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-469-880-6

```
Query Match          5.0%; Score 37.4; DB 2; Length 9323;
Best Local Similarity 48.0%; Pred. No. 0.38;
Matches 107; Conservative 0; Mismatches 116; Indels 0; Gaps
```

QY	498	GTTTTGTGAGACACTGGCCCATGAATGCTCTCCAGAAATAGTCGGATTTGGTCAA	557
Db	5502	GTTTCAAGTGACACAGATAACTGGAAGCCTTTTACTAAAGACGCTTCTCGCATCAAT	5443
QY	558	TAAATTTGGCAATCTACAGAACATGTGCTCTTTTCAGAGATTTATTTTAAATTAAC	617
Db	5442	GATCAGCAAGCCAGGCCAAGCAATCTGATATTGCACTCAGTTTAACTCGCTTTTATA	5383
QY	618	TATTTAAATAATTAAACATGTTACAAATTGCAATATAGTGAAGTGCCAAATCTCGCTAC	677
Db	5382	TTTTTCAATAATAAACCTGTTAAATAACTTTATTGTATCTCCGTAATTTTTTACAAG	5323
QY	678	ATGSGCTCAATCAGTTTTTACATATATTTCCACCCCATGTAATCA	720
Db	5322	ATGGGTAAGCCCTACTGAACTTTAAACCCACACAGTAATAA	5280

Search completed: April 11, 2003, 22:43:55  
Job time : 137.71 secs





US-09-764-872-639

Sequence 639, Application US/09764872

Publication No. US20030050231A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PAL25

CURRENT APPLICATION NUMBER: US/09/764,872

PRIOR FILING DATE: 2001-01-17

Prior application data removed - consult PALM or file wrapper

NUMBER OF SEQ ID NOS: 957

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 639

LENGTH: 3619

TYPE: DNA

ORGANISM: Homo sapiens

US-09-764-872-639

Query Match 6.1%; Score 45.4; DB 9; Length 3619;

Best Local Similarity 63.1%; Pred. No. 0.058;

Matches 70; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 617 TTATTAAAAATATTACATGGTACAAATTCATATAGTGAAGTGTGCAAACTTCGCTA 676

Db 560 TTGTGTTAGCAATTTTGATAAGATACAGCTTCAGACAGTAATACGCACACATCTTCAGT 619

QY 677 CATGGCTCAATGAGTTTTCATATATTTCCACCCATGTAATCACACCGA 727

Db 620 CAAAAGTCAATGAATGTTGACATATGTGTACCCCTGTGTATCATCTACTGA 670

RESULT 3

US-09-864-761-8570

Sequence 8570, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aemica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00671

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00672

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00673

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00674

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00675

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00676

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00677

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00678

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00679

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00680

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00681

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00682

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00683

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00684

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00685

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00686

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00687

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00688

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00689

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00690

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00691

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00692

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00693

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00694

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00695

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00696

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00697

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00698

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00699

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00700

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00701

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00702

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00703

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00704

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00705

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00706

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00707

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00708

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00709

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00710

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00711

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00712

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00713

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00714

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00715

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00716

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00717

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00718

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00719

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00720

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00721

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00722

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00723

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00724

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00725

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00726

PRIOR FILING DATE: 2001-01-

Db 411856 TTGATCTTAAGTATACACCCCAATGAATTTTACATATGTGTACACTAGTATACCTACCA 411915

Qy 724 CCGAGATCTAGA 735

Db 411916 CTTGAACCAAGA 411927

RESULT 5

US-09-864-761-8836/c  
; Sequence 8836, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: AecmIca-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 8836  
; LENGTH: 588  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO ALL159191.1  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 9.6  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 14  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 16  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.7  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 11  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 12

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 9.1  
US-09-864-761-8836

Query Match 5.7%; Score 42.8; DB 10; Length 588;  
Best Local Similarity 68.6%; Pred. No. 0.1;  
Matches 59; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 654 GTGAAGTGTGCAAACTCTCGCTACATGGCTCAATGAGTTTTTACATATATTTCCACCCAT 713  
Db 512 GTAAAGTGCACAGGCTCGAAATATACAGCTCAATAAATGTTTACATATATATGCGCCCAT 453  
Qy 714 GTAATCACCACCGAGATCTAGATAAG 739  
Db 452 ATAACCATCATTCAGATCAAGATGAG 427

RESULT 6

US-09-811-284-44  
; Sequence 44, Application US/09811284  
; Patent No. US20020058306A1  
; GENERAL INFORMATION:  
; APPLICANT: Vogeli, Gabriel  
; TITLE OF INVENTION: No. US20020058306A1 G Protein-Coupled Receptors  
; FILE REFERENCE: 00167US1  
; CURRENT APPLICATION NUMBER: US/09/811,284  
; CURRENT FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/189,783  
; PRIOR FILING DATE: 2000-03-16  
; PRIOR APPLICATION NUMBER: 60/189,907  
; PRIOR FILING DATE: 2000-03-16  
; PRIOR APPLICATION NUMBER: 60/189,918  
; PRIOR FILING DATE: 2000-03-16  
; PRIOR APPLICATION NUMBER: 60/189,960  
; PRIOR FILING DATE: 2000-03-16  
; PRIOR APPLICATION NUMBER: 60/189,917  
; PRIOR FILING DATE: 2000-03-16  
; PRIOR APPLICATION NUMBER: 60/192,945  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/192,916  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/192,923  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/192,933  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/192,830  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/192,234  
; PRIOR FILING DATE: 2000-03-27  
; PRIOR APPLICATION NUMBER: 60/192,155  
; PRIOR FILING DATE: 2000-03-24  
; PRIOR APPLICATION NUMBER: 60/192,935  
; PRIOR FILING DATE: 2000-03-29  
; NUMBER OF SEQ ID NOS: 258  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 44  
; LENGTH: 459  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-811-284-44

Query Match 5.7%; Score 42.6; DB 10; Length 459;  
Best Local Similarity 56.9%; Pred. No. 0.1;  
Matches 78; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Qy 599 ATTATTTTAACTTAATTAATTTTAAATAATTAAATGATGTAACATGTAACATTTGCATATAGTAA 658  
Db 107 ATTACCTTTTATTTATGTTTTCCTTTTTCAGTTTCCTTTTATGATAGAAAATTTACATGCGATAA 166  
Qy 659 GTGTGCAAACTCTCGCTACATGGCTCAATGAGTTTTTACATATATTTCCACCCATGTAAT 718  
Db 167 ATGCAGAACTTCAGTATACAAATTCATTTTGGTAAATATGTACACCTAGGCAAT 226  
Qy 719 CACCACCGAGATCTAGA 735

Db 227 TGACATCTCAATCAAGA 243

RESULT 7

US-09-901-136-3

; Sequence 3, Application US/09901136

; Publication No. US20030039988A1

; GENERAL INFORMATION:

; APPLICANT: HU, Song et al.

; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND

; TITLE OF INVENTION: USES THEREOF

; FILE REFERENCE: CL001273

; CURRENT APPLICATION NUMBER: US/09/901.136

; CURRENT FILING DATE: 2001-07-10

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 378361

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)...(378361)

; OTHER INFORMATION: n = A,T,C or G

US-09-901-136-3

Query Match 5.6%; Score 41.8; DB 9; Length 378361;

Best Local Similarity 51.9%; Pred. No. 7.3;

Matches 94; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 157 AVAAATTTGGCAATCTCAGAACATGTGCTTTTTCAGAGATTTATTTTAAATTAAC 616

Db 29238 AGAAGTTATGTCCTCAAAATTTACATATCCCTATGAAGTTATACCTGCTCTACTGTG 29297

QY 617 TTATTTAAAAATTAACATGGTACAAATTTGCATATATAGTGAAGTGTGCAAAATCTCGCTA 676

Db 29298 TTTATATCTACTTCTACTGAGGTATATTAACATGCAATTAATGCAAAATCTTGAGTG 29357

QY 677 CATGGCTCAATGATTTTACATATATTTCCACCCTGTAATCAACCCGAGATCTAGAA 736

Db 29358 CTTAGTTTGATGATTTAGATAATTTGTATACATCAACCGCCACCCCAAAACAAGAT 29417

QY 737 T 737

Db 29418 T 29418

RESULT 8

US-09-764-860-997

; Sequence 997, Application US/09764860

; Patent No. US20020094953A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC008

; CURRENT APPLICATION NUMBER: US/09/764,860

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 1198

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 997

; LENGTH: 290

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-764-860-997

Query Match 5.5%; Score 41.4; DB 10; Length 290;

Best Local Similarity 63.6%; Pred. No. 0.17;

Matches 63; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 637 GGTACAAATTTGCATATAGTGAAGTGTGCAAAATCTTCCTACATGGCTCAATGAGTTTTTA 696

Db 18 GGTAAATTTACATACAGTGAATGCACACATCATAGTCTACTGTTGTGCGTTTTGA 77

QY 697 CATATATTTCCACCATGTATATCACCACCGAGATCTAGA 735

Db 78 CAAATGCTTACACCTTTGTATTAATCAAAACCCCTTTTCAATA 116

RESULT 9

US-10-091-504-2274

; Sequence 2274, Application US/10091504

; Publication No. US20030059908A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC007C1

; CURRENT APPLICATION NUMBER: US/10/091,504

; CURRENT FILING DATE: 2002-03-07

; NUMBER OF SEQ ID NOS: 2442

; Prior Application removed - See File Wrapper or Palm

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2274

; LENGTH: 1402

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-091-504-2274

Query Match 5.5%; Score 41.2; DB 9; Length 1402;

Best Local Similarity 64.9%; Pred. No. 0.46;

Matches 61; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 637 GGTACAAATTTGCATATAGTGAAGTGTGCAAAATCTTCGCTACATGGCTCAATGAGTTTTTA 696

Db 872 GATATAACTAGTATTCAGTAAAGTGCACACGCTCTTAATTAAGGCTTCATGAATTATTA 931

QY 697 CATATATTTCCACCATGTATATCACCACCGAGAT 730

Db 932 TATTCAATATACACCATGTATCTACCTCCCTCAT 965

RESULT 10

US-10-091-504-2275

; Sequence 2275, Application US/10091504

; Publication No. US20030059908A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC007C1

; CURRENT APPLICATION NUMBER: US/10/091,504

; CURRENT FILING DATE: 2002-03-07

; NUMBER OF SEQ ID NOS: 2442

; Prior Application removed - See File Wrapper or Palm

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2275

; LENGTH: 1402

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-091-504-2275

Query Match 5.5%; Score 41.2; DB 9; Length 1402;

Best Local Similarity 64.9%; Pred. No. 0.46;

Matches 61; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 637 GGTACAAATTTGCATATAGTGAAGTGTGCAAAATCTTCGCTACATGGCTCAATGAGTTTTTA 696

Db 872 GATATAACTAGTATTCAGTAAAGTGCACACGCTCTTAATTAAGGCTTCATGAATTATTA 931

QY 697 CATATATTTCCACCATGTATATCACCACCGAGAT 730

Db 932 TATTCAATATACACCATGTATCTACCTCCCTCAT 965

RESULT 11

```
US-10-091-504-2276
; Sequence 2276, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2276
; LENGTH: 1402
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-504-2276

Query Match          5.5%; Score 41.2; DB 9; Length 1402;
Best Local Similarity 64.9%; Pred. No. 0.46;
Matches 61; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 637 GGTACAATTTGCATATAGTGAAGTGTGCAATCTTCGTACATGGCTCAATGAGTTTTTA 696
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 872 GATATAACTAGTATTTCAGTAAAGTGCACACGCTCTTAATTATAAGGCTTCATGAATTATTA 931
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 697 CATATATTTCCACCCATGTAATCACCACCGAGAT 730
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 932 TATTCATATACACCCATGTATCTACCTCCCTCAT 965
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
US-09-764-869-2274
; Sequence 2274, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2274
; LENGTH: 1402
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-2274

Query Match          5.5%; Score 41.2; DB 10; Length 1402;
Best Local Similarity 64.9%; Pred. No. 0.46;
Matches 61; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 637 GGTACAATTTGCATATAGTGAAGTGTGCAATCTTCGTACATGGCTCAATGAGTTTTTA 696
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 872 GATATAACTAGTATTTCAGTAAAGTGCACACGCTCTTAATTATAAGGCTTCATGAATTATTA 931
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 697 CATATATTTCCACCCATGTAATCACCACCGAGAT 730
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 932 TATTCATATACACCCATGTATCTACCTCCCTCAT 965
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-09-764-869-2275
; Sequence 2275, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
```

```
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2275
; LENGTH: 1402
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-2275

Query Match          5.5%; Score 41.2; DB 10; Length 1402;
Best Local Similarity 64.9%; Pred. No. 0.46;
Matches 61; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 637 GGTACAATTTGCATATAGTGAAGTGTGCAATCTTCGTACATGGCTCAATGAGTTTTTA 696
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 872 GATATAACTAGTATTTCAGTAAAGTGCACACGCTCTTAATTATAAGGCTTCATGAATTATTA 931
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 697 CATATATTTCCACCCATGTAATCACCACCGAGAT 730
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 932 TATTCATATACACCCATGTATCTACCTCCCTCAT 965
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
US-09-764-869-2276
; Sequence 2276, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2276
; LENGTH: 1402
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-2276

Query Match          5.5%; Score 41.2; DB 10; Length 1402;
Best Local Similarity 64.9%; Pred. No. 0.46;
Matches 61; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 637 GGTACAATTTGCATATAGTGAAGTGTGCAATCTTCGTACATGGCTCAATGAGTTTTTA 696
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 872 GATATAACTAGTATTTCAGTAAAGTGCACACGCTCTTAATTATAAGGCTTCATGAATTATTA 931
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 697 CATATATTTCCACCCATGTAATCACCACCGAGAT 730
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 932 TATTCATATACACCCATGTATCTACCTCCCTCAT 965
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
US-09-954-456-505
; Sequence 505, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
```

```
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 505
; LENGTH: 3777
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-505

Query Match      5.4%; Score 40.6; DB 10; Length 3777;
Best Local Similarity 55.2%; Pred. No. 1.2;
Matches 79; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

Qy 592 TTCAGAGATTATTTTAACTTAACTTATTTAAATATTAAACATGGTACAAATTGGCATA 651
Db 3303 TTCCAGCTGATGATCTAATATTTTAAACAAACATTATAGAGGTGTAATTATTTA 3362

Qy 652 TAGTGAAGTGTGCAATCTCGCTACATGGCTCAATGAGTTTACATATATTTCCACCC 711
Db 3363 CAATAAATGTCCTACTTTAAATATACAATTCAGTGAGTTTGCATAAATGATATACCC 3422

Qy 712 ATGTAATCACCACCGAGATCTAG 734
Db 3423 ATGTAACCAACACTCCAGTCAAG 3445
```

Search completed: April 12, 2003, 03:49:48  
Job time : 1009.98 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 14:22:18 ; Search time 1366.46 Seconds  
(without alignments)  
15994.713 Million cell updates/sec

Title: US-09-914-152-3\_COPY\_2000\_2750  
Perfect score: 751  
Sequence: 1 cctgccttggtcccaag.....tgatggaggctggaagttcc 751

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vi:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_inv:\*
- 32: em\_htg\_other:\*
- 33: em\_htg\_mus:\*
- 34: em\_htg\_pln:\*
- 35: em\_htg\_rod:\*
- 36: em\_htg\_mam:\*
- 37: em\_htg\_vit:\*
- 38: em\_sy:\*
- 39: em\_htgo\_hum:\*
- 40: em\_htgo\_mus:\*
- 41: em\_htgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	751	100.0	10562	6	E38420	Novel polyp
2	749.4	99.8	170121	9	AF064860	AF064860 Homo sapi
3	749.4	99.8	185982	2	AC073231	AC073231 Homo sapi
4	749.4	99.8	340000	9	HS21C080	AL163280 Homo sapi
5	420.4	56.0	2418	9	HS049973	U99773 Human Tigge
6	368.6	49.1	39444	9	AF036938	AF036938 Homo sapi
7	368.6	49.1	187520	2	AF282854	AF282854 Homo sapi
8	362.4	48.3	100418	9	HS1026E2	AL022143 Human DNA
9	362.4	48.3	158149	9	AC069000	AC069000 Homo sapi
10	360.8	48.0	138364	2	AC123789	AC123789 Homo sapi
11	354.2	47.2	115644	2	AC124056	AC124056 Homo sapi
12	354.2	47.2	115958	9	AC004736	AC004736 Human Chr
13	354	47.1	161339	9	AC007535	AC007535 Homo sapi
14	352.4	46.9	100020	9	AC092165	AC092165 Homo sapi
15	352.4	46.9	158276	2	AC012033	AC012033 Homo sapi
16	352.4	46.9	179816	2	AC074188	AC074188 Homo sapi
17	352.4	46.9	210301	9	AF307337	AF307337 Homo sapi
18	351.6	46.8	180519	2	AC024931	AC024931 Homo sapi
19	350.6	46.7	99342	9	HS262B17	AL008709 Human DNA
20	350.6	46.7	174380	2	HS201D17	AL020990 Human DNA
21	350.4	46.7	173559	9	AC016075	AC016075 Homo sapi
22	350.4	46.7	175559	9	AC112498	AC112498 Homo sapi
23	350	46.6	158958	9	AC007611	AC007611 Homo sapi
24	349.2	46.5	211662	9	AL391825	AL391825 Human DNA
25	348.6	46.4	172195	9	AC093862	AC093862 Homo sapi
26	348.2	46.4	153903	2	AC079836	AC079836 Homo sapi
27	348.2	46.4	199134	2	AC023458	AC023458 Homo sapi
28	348.2	46.4	208177	9	AC008072	AC008072 Homo sapi
29	347.6	46.3	51343	9	AC108017	AC108017 Homo sapi
30	347.4	46.3	161078	9	AP003500	AP003500 Homo sapi
31	347.4	46.3	161286	2	AC025120	AC025120 Homo sapi
32	346.4	46.1	60169	9	AL596257	AL596257 Human DNA
33	346	46.1	119220	9	AC016152	AC016152 Homo sapi
34	346	46.1	186851	9	AC007297	AC007297 Homo sapi
35	345.8	46.0	157122	2	AC026235	AC026235 Homo sapi
36	345.8	46.0	162638	2	AC096537	AC096537 Homo sapi
37	345	45.9	159098	2	AC016484	AC016484 Homo sapi
38	344.6	45.9	128800	9	AB053170	AB053170 Homo sapi
39	344.6	45.9	147798	9	AL390058	AL390058 Human DNA
40	344.6	45.9	165536	9	AC093600	AC093600 Homo sapi
41	344.6	45.9	179619	2	AC018596	AC018596 Homo sapi
42	344.4	45.9	119386	9	AC079824	AC079824 Homo sapi
43	344.4	45.9	140040	9	AC022029	AC022029 Homo sapi
44	344.4	45.9	166478	9	AC105750	AC105750 Homo sapi
45	343.4	45.7	180388	9	AC006206	AC006206 Homo sapi

ALIGNMENTS

RESULT 1  
E38420  
LOCUS E38420 Novel polypeptide.  
DEFINITION E38420  
ACCESSION E38420  
VERSION E38420.1 GI:18626994  
KEYWORDS JP 2000245464-A/2.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 10562)  
AUTHORS Narimatsu,H., Isshiki,S., Togayauchi,A. and Sasaki,K.  
TITLE Novel polypeptide  
JOURNAL Patent: JP 2000245464-A 2 12-SEP-2000;

E38420 10562 bp DNA linear PAT 31-JAN-2002

Pred. No. is the number of results predicted by chance to have a

COMMENT KYOWA HAKKO KOGYO CO LTD  
OS Homo sapiens (human)  
PW JP 2000245464-A/2  
PD 12-SEP-2000  
PR 25-FEB-1999 JP 1999047571  
PI HISASHI NARIMATSU, SOICHIRO ISSHIKI, AKIRA TOGAYAUCHI, PI  
KATSUTOSHI SASAKI  
PC C12N15/09.A01K67/027.C12N1/21.C12N5/10.C12N9/10.C12P19/00.PC  
C12P21/02.  
PC C12P21/08.C12Q1/68.G01N33/53/(C12N1/21.C12R1/185).(C12N5/10.  
C12R1/91).  
PC C12P21/02.C12R1/185).(C12P21/02.C12R1/91).C12N15/00.C12N5/00.  
PC (C12N5/00.C12R1/91)  
CC  
FH Key Location/Qualifiers  
FT promoter (1). (5000)  
FT exon (5001). (5140)  
FT exon (5001). (5273)  
FT exon (5459). (5567)  
FT exon (7427). (7586)  
FT exon (8234). (10562).  
FEATURES  
Source 1..10562  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 2610 a 2415 c 2574 g 2963 t  
ORIGIN  
Query Match 100.0%; Score 751; DB 6; Length 10562;  
Best Local Similarity 100.0%; Pred. No. 1e-145;  
Matches 751; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCTGCTTGGGCTCCAAAGTGTAGGATTACAGGGATGAGCCACCATGGTGGACTTCAT 60  
DB 2000 CCTGCTTGGGCTCCAAAGTGTAGGATTACAGGGATGAGCCACCATGGTGGACTTCAT 2059  
QY 61 GATAAATTCAGTGATGAGGAGCTCCCTCTTATGATGAACAAGAGGTTCTTG 120  
DB 2060 GATAAATTCAGTGATGAGGAGCTCCCTCTTATGATGAACAAGAGGTTCTTG 2119  
QY 121 AAATGAATCTACTCTGTGAGATGCTGTGAACATGTTGAATGACAGAAGAAT 180  
DB 2120 AAATGAATCTACTCTGTGAGATGCTGTGAACATGTTGAATGACAGAAGAAT 2179  
QY 181 TACAGTGTACATAGAGTTAGTTGATGAAGCAGTACAGGATTCGAGAGGATCGATTCCA 240  
DB 2180 TACAGTGTACATAGAGTTAGTTGATGAAGCAGTACAGGATTCGAGAGGATCGATTCCA 2239  
QY 241 ATTCAAAATAAGTTCTTCTGTGGTAAATGCTATCAAAATGGGTGCGATGCTACAGAG 300  
DB 2240 ATTCAAAATAAGTTCTTCTGTGGTAAATGCTATCAAAATGGGTGCGATGCTACAGAG 2299  
QY 301 AAATCTATCATGAAGAGAGTCAATTGATGTGGCAAACTTCATTGTTGCGTATTTTA 360  
DB 2300 AAATCTATCATGAAGAGAGTCAATTGATGTGGCAAACTTCATTGTTGCGTATTTTA 2359  
QY 361 AGAATTTGTCAGACACCCACCTTCACACCATGACCCCTGATCAGTCAGGAGCCAT 420  
DB 2360 AGAATTTGTCAGACACCCACCTTCACACCATGACCCCTGATCAGTCAGGAGCCAT 2419  
QY 421 CCACATTTAGGCGAGACCTCCAGCAGTAAAGATTATGATCTCTAAAGATCAGATG 480  
DB 2420 CCACATTTAGGCGAGACCTCCAGCAGTAAAGATTATGATCTCTAAAGATCAGATG 2479  
QY 481 AACATTAGCATTTTTTAAGCAATAAGTATTTTACGTAGATATGATGTTATTTTTT 540  
DB 2480 AACATTAGCATTTTTTAAGCAATAAGTATTTTACGTAGATATGATGTTATTTTTT 2539  
QY 541 AGGCATAATGCTATTATGATTTAATAGATCCAGTATATTGTAACATAACTTTAAATG 600  
DB 2540 AGGCATAATGCTATTATGATTTAATAGATCCAGTATATTGTAACATAACTTTAAATG 2599

QY 601 CACTGGGAGATAAAAGTATTGCTCTTTTATGATATTGCTTTTATTCAGTACTGCTGTA 660  
DB 2600 CACTGGGAGATAAAAGTATTGCTCTTTTATGATATTGCTTTTATTCAGTACTGCTGTA 2659  
QY 661 TGGAACTACATTTATCTCTTGGGTACACCTGTATACAGAAGAAATTTATCATGAGAAA 720  
DB 2660 TGGAACTACATTTATCTCTTGGGTACACCTGTATACAGAAGAAATTTATCATGAGAAA 2719  
QY 721 TGCTCATGCAATGATGGAGGCTGGAAGTCC 751  
DB 2720 TGCTCATGCAATGATGGAGGCTGGAAGTCC 2750  
RESULT 2  
AF064860 170121 bp DNA linear PRI 05-MAR-2002  
LOCUS Homo sapiens chromosome 21 clone PAC 70124 map 21q22.3, complete  
DEFINITION sequence.  
ACCESSION AF064860  
VERSION AF064860.2 GI:18958624  
KEYWORDS HTG; HTGS, DRAFT.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 170121)  
Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,  
Park,H.-S., Toyoda,A., Ishii,K., Toki,Y., Choh,D.-K., Soeda,E.,  
Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,  
Pollay,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,  
Patterson,D., Reichwald,K., Rump,A., Schillhabel,M.B., Schudy,A.,  
Zimmermann,W., Rosenthal,A., Kudoh,J., Kawasaki,K., Asakawa,S.,  
Shintani,A., Sasaki,T., Nagamine,K., Mitsuyama,S.,  
Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G.,  
Hornischer,K., Brandt,P., Scharfe,M., Schoen,O., Desario,A.,  
Reichelt,J., Kauer,G., Bloecker,H., Ramser,J., Beck,A., Klages,S.,  
Hennig,S., Riesselmann,L., Dagand,E., Haaf,T., Wehrmeyer,S.,  
Bonzym,K., Gardiner,K., Nizetic,D., Francis,F., Lehrach,H.,  
Reinhardt,R. and Yaspo,M.Laure.  
The DNA sequence of human chromosome 21  
Nature 405 (6784), 311-319 (2000)  
20289799  
10830953  
REFERENCE 2 (bases 1 to 170121)  
Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B.,  
Weber,J., Schattevoy,R., Yaspo,M.-L. and Rosenthal,A.  
Direct Submission  
Submitted (12-MAY-1998) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany  
3 (bases 1 to 170121)  
Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B.,  
Weber,J., Schattevoy,R., Yaspo,M.-L., Rosenthal,A., Yaspo,M.-L. and  
Rosenthal,A.  
Direct Submission  
Submitted (27-FEB-2002) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany  
4 (bases 1 to 170121)  
Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B.,  
Weber,J., Schattevoy,R., Yaspo,M.-L. and Rosenthal,A.  
Direct Submission  
Submitted (05-MAR-2002) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany  
COMMENT On Feb 27, 2002 this sequence version replaced gi:3171153.  
FEATURES  
Location/Qualifiers  
1..170121  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.3"  
/clone="PAC 70124"  
BASE COUNT 45184 a 36756 c 37509 g 50672 t  
ORIGIN

Query Match 99.8%; Score 749.4; DB 9; Length 170121;  
Best Local Similarity 99.9%; Pred. No. 1.3e-145;  
Matches 750; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTGCTTGGCTCCCAAGTGTAGGATACAGGATGAGCCACCATGGTGGTCTTCAAT 60  
|||||  
Db 83814 CTTGCTTGGCTCCCAAGTGTAGGATACAGGATGAGCCACCATGGTGGTCTTCAAT 83873

QY 61 GATAAACTTCACTGGATGAGGAGTCCCTCTTTATGATGAACAAAGAGGTGTTCTTGT 120  
|||||  
Db 83874 GATAAACTTCACTGGATGAGGAGTCCCTCTTTATGATGAACAAAGAGGTGTTCTTGT 83933

QY 121 AATGAATCTACTCTGTTGAGATGCTGTGAACATGTTGTAATGACAGAAAT 180  
|||||  
Db 83934 AATGAATCTACTCTGTTGAGATGCTGTGAACATGTTGTAATGACAGAAAT 83993

QY 181 TACAGTGTACATAGATTAGTATGATGAAGCAGTAGCAGGATTCAGAGGATCGATTCCA 240  
|||||  
Db 83994 TACAGTGTACATAGATTAGTATGATGAAGCAGTAGCAGGATTCAGAGGATCGATTCCA 84053

QY 241 ATTCAAAATAAGTCTCTGTTGGTAAAATGCTATCAAAATGGGCTCGCATGCTACAGAG 300  
|||||  
Db 84054 ATTCAAAATAAGTCTCTGTTGGTAAAATGCTATCAAAATGGGCTCGCATGCTACAGAG 84113

QY 301 AATCTATCATGAAGAGAGTCAATTGATGTGGCAAACTTCATTGTGCTGATTTTA 360  
|||||  
Db 84114 AATCTATCATGAAGAGAGTCAATTGATGTGGCAAACTTCATTGTGCTGATTTTA 84173

QY 361 AGAAATGTTCAGGACCAACCCACCTTCAACACCATGACCTGATCAGTACAGGACCAT 420  
|||||  
Db 84174 AGAAATGTTCAGGACCAACCCACCTTCAACACCATGACCTGATCAGTACAGGACCAT 84233

QY 421 CCACATTGAGGCGAGACCTCCAGCAGTAAAGATTATGATCTCTAAAGATCAGATG 480  
|||||  
Db 84234 CCACATTGAGGCGAGACCTCCAGCAGTAAAGATTATGATCTCTAAAGATCAGATG 84293

QY 481 AACATTAGCATTTTAAAGCAATTAAGTATTTTACGTAAAGATATGATGTTATTTT 540  
|||||  
Db 84294 AACATTAGCATTTTAAAGCAATTAAGTATTTTACGTAAAGATATGATGTTATTTT 84353

QY 541 AGCATAATGCTATTATGATTAATAGACTCAGTATATGTAACATACTTTAAATG 600  
|||||  
Db 84354 AGCATAATGCTATTATGATTAATAGACTCAGTATATGTAACATACTTTAAATG 84413

QY 601 CACTGGAGATAAAGTATTGCTCTTTTATGATTTGCTTTTATGCTAGTAGTCTGTAA 660  
|||||  
Db 84414 CACTGGAGATAAAGTATTGCTCTTTTATGATTTGCTTTTATGCTAGTAGTCTGTAA 84473

QY 661 TGGAACTACATTATCTTGGGTACACCTGTATACAGAAAGAAATTTATCATGAGGAAA 720  
|||||  
Db 84474 TGGAACTACATTATCTTGGGTACACCTGTATACAGAAAGAAATTTATCATGAGGAAA 84533

QY 721 TGCTCATGCAATGATGAGGCTGGAAGTCC 751  
|||||  
Db 84534 TGCTCATGCAATGATGAGGCTGGAAGTCC 84564

RESULT 3  
AC073231  
LOCUS AC073231 185982 bp DNA linear HTG 17-JUN-2000  
DEFINITION Homo sapiens chromosome 17 clone RP11-729A4, WORKING DRAFT  
SEQUENCE, 25 unordered pieces.  
ACCESSION AC073231  
VERSION AC073231.1 GI:8440043  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 185982)  
AUTHORS Waterston,R.H.  
TITLE The sequence of Homo sapiens clone  
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 185982)  
AUTHORS Waterston,R.H.  
TITLE Direct Submission  
JOURNAL Submitted (10-JUN-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

COMMENT  
\*\*\*\*\* Genome Center \*\*\*\*\*  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 25 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1724: contig of 1724 bp in length  
\* 1824: gap of unknown length  
\* 1835: contig of 1147 bp in length  
\* 2972: gap of unknown length  
\* 3072: contig of 1589 bp in length  
\* 4661: gap of unknown length  
\* 5930: contig of 1170 bp in length  
\* 6030: gap of unknown length  
\* 8127: contig of 2097 bp in length  
\* 8227: gap of unknown length  
\* 10160: contig of 1933 bp in length  
\* 10260: gap of unknown length  
\* 13236: contig of 2976 bp in length  
\* 13337: gap of unknown length  
\* 16604: gap of unknown length  
\* 16703: gap of unknown length  
\* 16704: contig of 3981 bp in length  
\* 20684: gap of unknown length  
\* 20784: contig of 4051 bp in length  
\* 24835: gap of unknown length  
\* 24935: contig of 4792 bp in length  
\* 29221: gap of unknown length  
\* 29828: contig of 5120 bp in length  
\* 34947: gap of unknown length  
\* 35047: contig of 6117 bp in length  
\* 41164: gap of unknown length  
\* 41264: contig of 5786 bp in length  
\* 47050: gap of unknown length  
\* 47051: contig of 5934 bp in length  
\* 47151: gap of unknown length  
\* 53084: contig of 7440 bp in length  
\* 53185: gap of unknown length  
\* 60624: gap of unknown length  
\* 60725: contig of 8333 bp in length  
\* 69058: gap of unknown length  
\* 69157: contig of 9372 bp in length  
\* 78529: gap of unknown length  
\* 78629: contig of 6846 bp in length  
\* 85476: gap of unknown length  
\* 85575: contig of 11576 bp in length  
\* 97151: gap of unknown length  
\* 97252: contig of 15860 bp in length  
\* 113111: gap of unknown length  
\* 113212: contig of 16659 bp in length  
\* 129870: gap of unknown length  
\* 129871: contig of 15598 bp in length  
\* 145569: gap of unknown length  
\* 145669: contig of 18928 bp in length  
\* 164597: gap of unknown length  
\* 164597: contig of 21286 bp in length.

FEATURES  
source  
1. .185982  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/clone="RP11-729A4"  
1. .1724  
/note="assembly\_name:Contig3"

misc\_feature  
1. .1724

misc\_feature 1825.. 2971 /note="assembly\_name:Contig4"  
misc\_feature 3072.. 4660 /note="assembly\_name:Contig5"  
misc\_feature 4761.. 5930 /note="assembly\_name:Contig6  
clone\_end:17  
vector\_side:right  
misc\_feature 6031.. 8127 /note="assembly\_name:Contig7"  
misc\_feature 8228.. 10160 /note="assembly\_name:Contig8"  
misc\_feature 10261.. 13236 /note="assembly\_name:Contig9"  
misc\_feature 13337.. 16603 /note="assembly\_name:Contig10"  
misc\_feature 16704.. 20684 /note="assembly\_name:Contig11"  
misc\_feature 20785.. 24835 /note="assembly\_name:Contig12"  
misc\_feature 24936.. 29727 /note="assembly\_name:Contig13  
clone\_end:SP6  
vector\_side:right  
misc\_feature 29828.. 34947 /note="assembly\_name:Contig14"  
misc\_feature 35048.. 41164 /note="assembly\_name:Contig15"  
misc\_feature 41265.. 47050 /note="assembly\_name:Contig16"  
misc\_feature 47151.. 53084 /note="assembly\_name:Contig17"  
misc\_feature 53185.. 60624 /note="assembly\_name:Contig18"  
misc\_feature 60725.. 69057 /note="assembly\_name:Contig19"  
misc\_feature 69158.. 78529 /note="assembly\_name:Contig20"  
misc\_feature 78630.. 85475 /note="assembly\_name:Contig21"  
misc\_feature 85576.. 97151 /note="assembly\_name:Contig22"  
misc\_feature 97252.. 113111 /note="assembly\_name:Contig23"  
misc\_feature 113212.. 129870 /note="assembly\_name:Contig24"  
misc\_feature 129971.. 145568 /note="assembly\_name:Contig25"  
misc\_feature 145669.. 164596 /note="assembly\_name:Contig26"  
misc\_feature 164697.. 185982 /note="assembly\_name:Contig27"  
BASE COUNT 53704 a 39120 c 39561 g 51192 t 2405 others  
ORIGIN  
Query Match 99.8%; Score 749.4; DB 2; Length 185982;  
Best Local Similarity 99.9%; Pred. No. 1.3e-145;  
Matches 750; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CTGCGCTTGGGTCCTCCAAAGTCTAGGATACAGGATGAGCCACCATGTCGACTTCAT 60  
Db 26697 CTGCGCTTGGGTCCTCCAAAGTCTAGGATACAGGATGAGCCACCATGTCGACTTCAT 26756  
QY 61 GATAAATTCAGTGGATGAGGAGTGCCTCTTATGATGAACAAGAAGTGGTTCTTG 120  
Db 26757 GATAAATTCAGTGGATGAGGAGTGCCTCTTATGATGAACAAGAAGTGGTTCTTG 26816  
QY 121 AAATGGAATCTACTCCTGCTGAAGTCTGTGACATCTTGAATCACAAGAAGATT 180  
Db 26817 AAATGGAATCTACTCCTGCTGAAGTCTGTGACATCTTGAATCACAAGAAGATT 26876  
QY 181 TACAGTGTACATAGATTAGTTGATGAAGCAGTAGCAGGATTCGAGAGGATCGATTCCA 240

Db 26877 TACAGTGTACATAGATTAGTTGATGAAGCAGTAGCAGGATTCGAGAGGATCGATTCCA 26936  
QY 241 ATTTCAAATAAAGTTCTTCTGTGGTAAATGCTATCAAAATGGCGTCGCATGCTACAGAG 300  
Db 26937 ATTTCAAATAAAGTTCTTCTGTGGTAAATGCTATCAAAATGGCGTCGCATGCTACAGAG 26996  
QY 301 AAATCTATCATGAAGGAAGAGTCAATTTGATGTGGCAAACTTCATTTGTTGCTATTATTA 360  
Db 26997 AAATCTATCATGAAGGAAGAGTCAATTTGATGTGGCAAACTTCATTTGTTGCTATTATTA 27056  
QY 361 AGAAATTTGTCAGGACCCCAACCTTCAACAACCATGACCTGATCATCAGGAGCCAT 420  
Db 27057 AGAAATTTGTCAGGACCCCAACCTTCAACAACCATGACCTGATCATCAGGAGCCAT 27116  
QY 421 CCACATTGAGCGAGAACCTCCACGAGTAAAGAAATTTATGATTTCTCTAAAGGATCAGATG 480  
Db 27117 CCACATTGAGCGAGAACCTCCACGAGTAAAGAAATTTATGATTTCTCTAAAGGATCAGATG 27176  
QY 481 AACATTAGCATTTTTTTTAAAGCAATAAGTATTTTTTACGTAAAGATATGATGTTATTTTTT 540  
Db 27177 AACATTAGCATTTTTTTTAAAGCAATAAGTATTTTTTACGTAAAGATATGATGTTATTTTTT 27236  
QY 541 AGCATTAATGCTATTATGATTTAATAGACTCCAGTATTTTGTAAACATAACTTTTAAATG 600  
Db 27237 AGCATTAATGCTATTATGATTTAATAGACTCCAGTATTTTGTAAACATAACTTTTAAATG 27296  
QY 601 CACTGGGAGATAAAAGTATTTGCTCTTTTATGATATTTTGCCTTTATGCTAGTCTGTAA 660  
Db 27297 CACTGGGAGATAAAAGTATTTGCTCTTTTATGATATTTTGCCTTTATGCTAGTCTGTAA 27356  
QY 661 TGGAACTACATATCTCTGGGTACACCTGTATACAGAAAGAAATTTTATCATGAGGAAA 720  
Db 27357 TGGAACTACATATCTCTGGGTACACCTGTATACAGAAAGAAATTTTATCATGAGGAAA 27416  
QY 721 TGCTCATGCAATGATGGAGGCTGGAAGTCC 751  
Db 27417 TGCTCATGCAATGATGGAGGCTGGAAGTCC 27447  
RESULT 4  
HS21C080 340000 bp DNA linear PRI 24-MAY-2000  
LOCUS Homo sapiens chromosome 21 segment HS21C080.  
DEFINITION AL163280 AF001735 BA000005  
ACCESSION AL163280.2 GI:7717369  
VERSION human.  
KEYWORDS human.  
SOURCE  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 340000)  
AUTHORS Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,  
Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,  
Ohki,M., Takagi,T., Sakaki,Y., Kump,K., Taudien,S., Blechschmidt,K.,  
Polley,A., Menzel,U., Delabar,J., Kump,K., Lehmann,R.,  
Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,  
Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K.,  
Asakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuayama,S.,  
Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G.,  
Hornischer,K., Brandt,P., Scharfe,M., Schoen,O., Desario,A.,  
Reichelt,J., Kauer,G., Bloecker,H., Ramser,J., Beck,A., Klages,S.,  
Hennig,S., Riesselmann,L., Dagand,E., Wehrmeyer,S., Borzym,K.,  
Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and  
Vaspo,M.L.  
Direct Submission  
TITLE Submitted (05-MAY-2000) The Chromosome 21 Mapping and Sequencing  
JOURNAL Consortium: \* RIKEN Genomic Sciences Center, Human Genome Research  
Group \* Institute of Molecular Biotechnology, Genome Analysis \*  
Keio University School of Medicine, Dept. of Molecular Biology \*  
GBF, Dept. of Genome Analysis \* Max-Planck Institute for Molecular  
Genetics (addresses see below)  
The Chromosome 21 Mapping and Sequencing Consortium consists of  
\* RIKEN Genomic Sciences Center, Human Genome Research Group, \*

Sagamihara 228-8555, Japan,  
\* e.mail: sakaki@gsc.riken.go.jp  
\* URL: http://hgp.gsc.riken.go.jp/  
and  
\* Institute of Molecular Biotechnology, Genome Analysis, \*  
Beutenbergstrasse 11, D-07745 Jena, Germany,  
\* e.mail: gscj-submit@genome.imb-jena.de  
\* URL: http://genome.imb-jena.de/  
and  
\* Keio University School of Medicine, Dept. of Molecular Biology, \*  
Tokyo 160-8582, Japan,  
\* e.mail: shimizu@dbm-med.keio.ac.jp  
\* URL: http://adenine.dmb.med.keio.ac.jp/  
and  
\* GBF, Dept. of Genome Analysis,  
\* Mascheroder Weg 1, D-38124 Braunschweig, Germany, \* e.mail:  
info.genome@gbf.de  
\* URL: http://genome.gbf.de/  
and  
\* Max-Planck Institute for Molecular Genetics,  
\* Ihnestrasse 73, D-14195 Berlin, Germany,  
\* e.mail: info-chr21@molgen.mpg.de  
\* URL: http://chr21.rz-berlin.mpg.de/  
Location/Qualifiers  
1. 340000  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.3"  
c1. 125946  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.3"  
/clone="P160L9, 5' partial"  
/clone\_lib="RPC11.3-5 PAC library"  
/note="Accession No. AF121897"  
73410. 243533  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.3"  
/clone="P70124"  
/clone\_lib="RPC11.3-5 PAC library"  
/note="Accession No. AF064860"  
224137. >340000  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.3"  
/clone="P206A10, 3' partial"  
/clone\_lib="RPC11.3-5 PAC library"  
/note="Accession No. AF121782"  
286628. >340000  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.3"  
/clone="BAC-291B3, 3' partial"  
/clone\_lib="BAC library"  
/note="Accession No. AF064857"  
1088. 1566  
/note="L1M4"  
/rpt\_family="LINE/L1"  
/rpt\_type=DISPERSED  
2398. 2495  
/note="(TCTA)n"  
/rpt\_family="Simple\_repeat"  
/rpt\_type=TANDEM  
complement(2496. 2878)  
/note="THEIC"  
/rpt\_family="LTR/MaLR"  
/rpt\_type=DISPERSED

FEATURES

repeat\_region  
2879. 2931  
/note="(TCTA)n"  
/rpt\_family="Simple\_repeat"  
/rpt\_type=TANDEM  
2932. 3004  
/note="(TCCA)n"  
/rpt\_family="Simple\_repeat"  
/rpt\_type=TANDEM  
3005. 3074  
/note="(TCCA)n"  
/rpt\_family="Simple\_repeat"  
/rpt\_type=TANDEM  
3473. 19157  
/gene="SH3BGR"  
join(<3473. 3565,12649. 12678,15331. 15460,18739. 19157)  
/gene="SH3BGR"  
join(3473. 3565,12649. 12678,15331. 15426)  
/partial  
/gene="SH3BGR"  
/note="Accession No. X93498"  
/codon\_start=1  
/product="21-Glutamic Acid Rich protein 21-GARP"  
/protein\_id="CAB90445.1"  
/db\_xref="GI:7717370"  
/translation="GSEKAEGETEAQKGESEVDGNLPEAQKNEEGETATEETEE  
TAMEGAEGFAEEETAEGEPEDEDS"  
3473. 3565  
/gene="SH3BGR"  
/number=4  
5542. 5829  
/note="AluX"  
/rpt\_family="SINE/Alu"  
/rpt\_type=DISPERSED  
complement(6887. 7067)  
/note="L1MD3"  
/rpt\_family="LINE/L1"  
/rpt\_type=DISPERSED  
complement(8482. 8762)  
/note="AluJo"  
/rpt\_family="SINE/Alu"  
/rpt\_type=DISPERSED  
complement(8865. 8964)  
/note="L2"  
/rpt\_family="LINE/L2"  
/rpt\_type=DISPERSED  
9131. 9424  
/note="AluSq"  
/rpt\_family="SINE/Alu"  
/rpt\_type=DISPERSED  
complement(9721. 9882)  
/note="L2"  
/rpt\_family="LINE/L2"  
/rpt\_type=DISPERSED  
complement(9989. 10077)  
/note="L2"  
/rpt\_family="LINE/L2"  
/rpt\_type=DISPERSED  
complement(10078. 10429)  
/note="THEIB"  
/rpt\_family="LTR/MaLR"  
/rpt\_type=DISPERSED  
complement(10430. 10651)  
/note="L2"  
/rpt\_family="LINE/L2"  
/rpt\_type=DISPERSED  
10980. 11000  
/note="(TTTA)n"  
/rpt\_family="Simple\_repeat"  
/rpt\_type=TANDEM  
complement(11002. 11282)  
/note="AluSp"  
/rpt\_family="SINE/Alu"  
/rpt\_type=DISPERSED

Db 157882 TGGAAACTACATTATCTCTTTGGGTACACCTCTATACAGAAAGAAATTTATCATGAGGAAA 157941

Qy 721 TGCTCATGCAATGATGAGGCTGCAAGATGCC 751  
|||||  
Db 157942 TGCTCATGCAATGATGAGGCTGCAAGATGCC 157972

RESULT 5  
HSU49973 2418 bp DNA linear PRI 28-JUN-1997  
LOCUS Human fligerl transposable element, complete consensus sequence.  
DEFINITION U49973  
ACCESSION U49973.1 GI:2226003  
VERSION  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2418)  
Smit,A.F. and Riggs,A.D.  
Tigers and DNA transposon fossils in the human genome  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 93 (4), 1443-1448 (1996)  
MEDLINE 96202298  
PUBMED 8643651  
2 (bases 1 to 2418)  
Robertson,H.M.  
Members of the pogo superfamily of DNA-mediated transposons in the  
TITLE human genome  
JOURNAL Mol. Gen. Genet. 252 (6), 761-766 (1996)  
MEDLINE 97074895  
PUBMED 8917322  
3 (bases 1 to 2418)  
Robertson,H.M.  
Direct Submission  
TITLE Submitted (15-FEB-1996) Hugh M. Robertson, Entomology, University  
JOURNAL of Illinois at Urbana-Champaign, 505 S. Goodwin, Urbana, IL 61801,  
USA  
COMMENT On Jun 28, 1997 this sequence version replaced gi:1224064.  
FEATURES  
source Location/Qualifiers  
1..2418  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/transposon="Tiggerl"  
/note="consensus sequence based on 50 full-length genomic  
sequences"  
repeat\_region 1..13  
/rpt\_type=inverted  
CDS 425..1789  
/note="ORF1; MER37; putative transposase similar to pogo  
element"  
/citation=[1]  
/codon\_start=1  
/protein\_id="AAB61714.1"  
/db\_xref="GI:2226004"  
/translation="MASKCSERKSRKSTSLTNQLEMIKLSSEGMSKAEIGRKLGLLR  
QTGVVNAKKEFLKEIKSATPVNTRMIRKNSLIADMEKVLVVMIEDQSHNIPLSQ  
SLIOSKALTFLNSMKAEGEEAAEKLASRGWFMERKPSRLHNKVGQEAASADGE  
AAASYPEDLAKIIDGGVTKQIFNVDETFATYKKMPSRTFIAREKSMGFKASKDR  
LTLGLNAAQDFKLKPLMILYHSENPRLAKNYAKSTLPVLYKWNKAMWTAHLFTAM  
TEYFKPTVETYSERKKIFKILLIDNAEPHPRALMEMYKEINVVFPAANTTSILQPM  
DQGVISTFKSYVLRNTRFKAAIAAIDSDSGSQSKLTFWKGFITLDAIKNRDSDWE  
EVKISTLTGVNKKLIPTLMDDFEGKTSVEEYADVVEIARELELEVEPEDVTLLQS  
HDKN"  
CDS 1811..2206  
/note="ORF2: function unknown"  
/codon\_start=1  
/protein\_id="AAB61715.1"  
/db\_xref="GI:2226005"  
/translation="MDEORKWFLFEMESTGCEDAVNIVEMTTKOLEYINLVDKAAGF  
ERISNPERSTVQKMLNSIACYREIFRFSQSMRQTSLLSIFKRLPOPPPSAIT  
TLISQQPSTSRQDPPPKAKRLRLAEGSDDR"  
polyA\_signal 2218..2223  
repeat\_region 2405..2418

repeat_region	complement(11315..11556)
	/note="AluJo"
	/rpt_family="SINE/Alu"
	/rpt_type=DISPERSED
repeat_region	11702..11753
	/note="MIR"
	/rpt_family="SINE/MIR"
	/rpt_type=DISPERSED
exon	12649..12678
	/gene="SH3BGR"
	/number=5
repeat_region	complement(13115..13425)
	/note="AluX"
	/rpt_family="SINE/Alu"
	/rpt_type=DISPERSED
repeat_region	complement(13598..13892)
	/note="AluY"
	/rpt_family="SINE/Alu"
	/rpt_type=DISPERSED
repeat_region	14013..14131
	/note="(CA)n"
	/rpt_family="Simple_repeat"
Query Match 99.88; Score 749.4; DB 9; Length 340000; Best Local Similarity 99.94; Pred. No. 1.2e-145; Matches 750; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	1 CCTCGCTGGGCTCCCAAAGTGCTAGGATTACAGGATGAGCCACCATGGTCGACTTCAT 60
Db	157222 CCTCGCTGGGCTCCCAAAGTGCTAGGATTACAGGATGAGCCACCATGGTCGACTTCAT 157281
QY	61 GATAAACTTCAGTGGATGAGGAGTCGCCTTTATGATGAACAAAGAGGTGTTCTTG 120
Db	157282 GATAAACTTCAGTGGATGAGGAGTCGCCTTTATGATGAACAAAGAGGTGTTCTTG 157341
QY	121 AAATGGAATCTACTCCCTGGTGAAGATGCTGTGAACATTGTTGAAATGACAGAAGAAT 180
Db	157342 AAATGGAATCTACTCCCTGGTGAAGATGCTGTGAACATTGTTGAAATGACAGAAGAAT 157401
QY	181 TACAGTGTACATAGAGTTAGTTGATGAAGCAGTAGCAGGATTCGAGAGGATCGATTCCA 240
Db	157402 TACAGTGTACATAGAGTTAGTTGATGAAGCAGTAGCAGGATTCGAGAGGATCGATTCCA 157461
QY	241 ATTTCAAATAAGTCTCTCTGTGGGTAAATGCTATCAATGGCGTCGCATCGCATACAG 300
Db	157462 ATTTCAAATAAGTCTCTCTGTGGGTAAATGCTATCAATGGCGTCGCATCGCATACAG 157521
QY	301 AAATCTATCATGAAGAGAGAGTCAATGTGTCGCAACTTCATTGTTGCTATTATTTA 360
Db	157522 AAATCTATCATGAAGAGAGAGTCAATGTGTCGCAACTTCATTGTTGCTATTATTTA 157581
QY	361 AGAAATTGTCAGGACCAACCCCAACCTTCAACAACCATGACCCCTCATCAGTCAGGAGCAT 420
Db	157582 AGAAATTGTCAGGACCAACCCCAACCTTCAACAACCATGACCCCTCATCAGTCAGGAGCAT 157641
QY	421 CCACATTGAGCGGAGACCTCCACAGTAAAGAGATTATGATTCCTAAAGGATCAGATG 480
Db	157642 CCACATTGAGCGGAGACCTCCACAGTAAAGAGATTATGATTCCTAAAGGATCAGATG 157701
QY	481 AACATTAGCATTTTTTTTAAGCAATAAAGTATTTTTTACGTAGATATGTATGTTATTTTT 540
Db	157702 AACATTAGCATTTTTTTTAAGCAATAAAGTATTTTTTACGTAGATATGTATGTTATTTTT 157761
QY	541 AGGCATAATGCTATTATGCAATTAATAGACTCCAGTATATTGTAAGACATAACTTTAAATG 600
Db	157762 AGGCATAATGCTATTATGCAATTAATAGACTCCAGTATATTGTAAGACATAACTTTAAATG 157821
QY	601 CACTGGGAGATAAAGATTGTCCTTTTATGATATTGCTTTATTGCGTAGTCTGTAA 660
Db	157822 CACTGGGAGATAAAGATTGTCCTTTTATGATATTGCTTTATTGCGTAGTCTGTAA 157881
QY	661 TGGAACTACATTATCTCTTGGGTACACCTGTATACAGAAGAATTTATCATGAGAAA 720
Db	157882 TGGAACTACATTATCTCTTGGGTACACCTGTATACAGAAGAATTTATCATGAGAAA 157941

TITLE	Direct Submission
JOURNAL	Submitted (04-DBC-1997) Genome Analysis, Institut für Molecular Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
FEATURES	Location/Qualifiers
source	1..39444
repeat_region	/organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="X" /map="Xq28" /clone="cosmid Qc8C8" 6..59
repeat_region	/evidence-not_experimental /rpt_family="A" 60..198
repeat_region	/evidence-not_experimental /rpt_family="A" 222..428
repeat_region	/evidence-not_experimental /rpt_family="A" 429..481
repeat_region	/evidence-not_experimental /rpt_family="LlPA" 483..546
repeat_region	/evidence-not_experimental /rpt_family="LlPA" 547..951
repeat_region	/evidence-not_experimental /rpt_family="Llp" 982..1099
repeat_region	/evidence-not_experimental /rpt_family="LlPA" 1100..1254
repeat_region	/evidence-not_experimental /rpt_family="LlPA" 1254..1289
repeat_region	/evidence-not_experimental /rpt_family="LlPA" 1301..1348
repeat_region	/evidence-not_experimental /rpt_family="LlPA" 1349..1413
repeat_region	/evidence-not_experimental /rpt_family="Llp" 1414..1456
repeat_region	/evidence-not_experimental /rpt_family="LlPA" 1457..1519
repeat_region	/evidence-not_experimental /rpt_family="LlM" 1520..1567
repeat_region	/evidence-not_experimental /rpt_family="LlM" 1568..1643
repeat_region	/evidence-not_experimental /rpt_family="A" 1644..1706
repeat_region	/evidence-not_experimental /rpt_family="A" 1707..1868
misc_feature	/evidence-not_experimental /rpt_family="A" 1869..5848
repeat_region	/note="GC score = 7.60 (134bp); GC region" 5849..5985
repeat_region	/evidence-not_experimental /rpt_family="LlM" 5986..6217
repeat_region	/evidence-not_experimental /rpt_family="LlM" 6218..6217

repeat\_region 6222..6279  
/evidence-not\_experimental  
/rpt\_family="L1MA"  
repeat\_region 6296..6392  
/evidence-not\_experimental  
/rpt\_family="L1M"  
repeat\_region 6861..6882  
/note="l11 with 91% homology to l12"  
/evidence-not\_experimental  
/rpt\_type-inverted  
repeat\_region 6905..6926  
/note="l12 with 91% homology to l11"  
/evidence-not\_experimental  
/rpt\_type-inverted  
exon 9386..10151  
/note="GRAIL, score = 100.000%, comment = excellent"  
/evidence-not\_experimental  
exon 10100..10151  
/note="Xpound exon prediction, score = 87% (0%)"  
/evidence-not\_experimental  
repeat\_region 11540..11628  
/evidence-not\_experimental  
/rpt\_family="LIN"  
repeat\_region 12813..12925  
/evidence-not\_experimental  
/rpt\_family="A"  
repeat\_region 12935..12958  
/note="211 with 88% homology to 212"  
/evidence-not\_experimental  
/rpt\_type-inverted  
repeat\_region 12986..13009  
/note="212 with 88% homology to 211"  
/evidence-not\_experimental  
/rpt\_type-inverted  
repeat\_region 13031..13077  
/evidence-not\_experimental  
/rpt\_family="100"  
repeat\_region 13096..13170  
/evidence-not\_experimental  
/rpt\_family="100"  
repeat\_region 13171..13223  
/evidence-not\_experimental  
/rpt\_family="100"  
repeat\_region 13454..13489  
/note="homology = 88.90%, score = 24, counts = 9"  
/evidence-not\_experimental  
/rpt\_type-tandem  
repeat\_region 13509..13563  
/evidence-not\_experimental  
/rpt\_family="100"  
repeat\_region 13592..13769  
/evidence-not\_experimental  
/rpt\_family="100"  
repeat\_region 13813..13903  
/note="homology = 80.20%, score = 42, counts = 7"  
/evidence-not\_experimental  
/rpt\_type-tandem  
repeat\_region 14136..14660  
/evidence-not\_experimental  
/rpt\_family="Lip"  
repeat\_region 14753..14835  
/evidence-not\_experimental  
repeat\_region 14864..15065  
/evidence-not\_experimental  
repeat\_region 15066..15130  
/evidence-not\_experimental  
repeat\_region 15131..15249  
/evidence-not\_experimental  
repeat\_region 15250..15514  
/evidence-not\_experimental  
repeat\_region 15517..15544

/evidence-not\_experimental  
15552..15614  
/evidence-not\_experimental  
15618..15755  
/evidence-not\_experimental  
15756..15815  
/evidence-not\_experimental  
15850..15899  
/evidence-not\_experimental  
15901..16008  
/evidence-not\_experimental  
16011..16124  
/evidence-not\_experimental  
16158..16210  
/evidence-not\_experimental  
16211..16271  
/evidence-not\_experimental  
16275..16379  
/evidence-not\_experimental  
16380..16501  
/evidence-not\_experimental  
16506..16799  
/evidence-not\_experimental  
16801..17012  
/evidence-not\_experimental  
17056..17097  
/evidence-not\_experimental  
/rpt\_family="L1PA"  
repeat\_region 17099..17161  
/evidence-not\_experimental  
/rpt\_family="L1PA"  
repeat\_region 17167..17317  
/evidence-not\_experimental  
/rpt\_family="L1PA"  
repeat\_region 17318..17361  
/evidence-not\_experimental

Query Match 49.1%; Score 368.6; DB 9; Length 39444;  
Best Local Similarity 76.3%; Pred. No. 9.9e-67;  
Matches 495; Conservative 0; Mismatches 144; Indels 10; Gaps 3;  
QY 52 CGACTTCATGATAAACTTCAGTGGATGAGGAGCTGCTCTTATGATGAACAAAGAGGT 111  
DB 25325 CAATCTCGTGAATAAACTTGAAGGATGAGTGTCTTCTTACAGACGAGCATGAGAGT 25382  
QY 112 GGTTCTTTGAATGGAATCTACTCTCTGTTGAGAGATGCTGTGAAATTTGTTGAATGACAA 171  
DB 25383 GATTTCTTGATAAGGAATCTACTCTAGTAAAGATGCTGTGAATATTGTTGAATGACAA 25442  
QY 172 GAAAGAAATTTACAGTGTACATAGAGTTAGTTGATGAAGCAGTAGCAGGATTCGAGAGGA 231  
DB 25443 CAAAGGATTTAGAAATATTACATAAACTTAGTTGATGAAGCAGTGCAGGGTTTAAGAGGA 25502  
QY 232 TCGATTCCAAATTTCAAATAAGTTCTTCTGTGGGTAAATGCTATCAAAATGGCGTCGCAT 291  
DB 25503 TTGGCTCCATTTTGAAGAAGAGTTCTACTGTGGGTAAATGCGCAACACATCATCAT 25562  
QY 292 GCTACAGAGAAATCTATCATGAAAAGAGAGTCAATTTGATGTGGCAAACTTCATTTGTTGT 351  
DB 25563 GCTGAAGATAAATCTTTTGTGAAAAGAGAGTCAATTTGATGAGGCAAACTTCATTTGTTGT 25622  
QY 352 CGTATTTTAAGAATTTGTGAGGACCCCAACCTTCAACACCATGACCCCTGATCAGTC 411  
DB 25623 CTTATTTTAAGAAATTTGCCACAGCTGCCCAACCTTTAGCAACCACTGCCCTGACTGGTC 25682  
QY 412 AGGAGCCATCCACATTCAGGGGGAACCTCCAGCAGTAAAAGATATGATTTCTCTAAAG 471  
DB 25683 AGCAGCCATCCACATTCAGGCAAACTTCCACCAGGAAGATTTATGATTTGTTGAAG 25742  
QY 472 GATCAGATGAACATTTAGCATTTTTTTTAAAGCAATAAGATTTTTTACGTAAGATATGTATG 531  
DB 25743 CGTCAGATGATTTCTAGCATTTTTTTT-AGCAATAAGATTTTTTTAATTAAGGATATGTATAA 25801





sequence is ambiguous, there is an annotation using the 'unsure' feature key.  
This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr1>  
102562 is from the library RPC15 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/VECTOR:pcYPAC2>.

FEATURES

Location/Qualifiers  
1. .100418  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/map="q24.1-25.3"  
/clone="RP5-1026E2"  
/clone\_lib="RPC1-5"  
9. .1897  
/note="Retroviral Sequence"  
repeat\_region  
2018. .2177  
/note="2 copies 80 mer 95% conserved"  
prim\_transcript  
2052. .>2289  
/note="match: EST AA601759 clone IMAGE:1099629"  
repeat\_region  
2176. .2339  
/note="2 copies 82 mer 85% conserved"  
prim\_transcript  
complement(2423. .3021)  
/note="match: 3' EST AA783035 clone J375023"  
repeat\_region  
complement(3220. .3367)  
/note="pIR5 repeat: matches 288. .142 of consensus"  
repeat\_region  
complement(3265. .3898)  
/note="pIR5 repeat: matches 2427. .1800 of consensus"  
repeat\_region  
complement(3900. .3945)  
/note="pIR5 repeat: matches 1397. .1353 of consensus"  
misc\_feature  
3987. .4286  
/note="match: 253809 STS containing (CA) repeat"  
repeat\_region  
4102. .4127  
/note="13 copies of AC 100% conserved; differs from 253809"  
repeat\_region  
5110. .5398  
/note="AluJb repeat: matches 12. .302 of consensus"  
repeat\_region  
complement(5436. .5576)  
/note="MIR2 repeat: matches 146. .4 of consensus"  
repeat\_region  
5506. .5937  
/note="MER2 repeat: matches 8. .344 of consensus"  
misc\_feature  
complement(6655. .7102)  
/note="match: GSS B82054"  
repeat\_region  
complement(7192. .7423)  
/note="L1A10 repeat: matches 900. .662 of consensus"  
repeat\_region  
9003. .9267  
/note="MIR repeat: matches 2. .262 of consensus"  
repeat\_region  
complement(13451. .13902)  
/note="MLT1D repeat: matches 505. .44 of consensus"  
repeat\_region  
14327. .14444  
/note="MIR repeat: matches 47. .164 of consensus"  
repeat\_region  
complement(14857. .14993)  
/note="MER5A repeat: matches 179. .37 of consensus"  
repeat\_region  
15170. .15209  
/note="10 copies 4 mer tcat 83% conserved"  
repeat\_region  
complement(19889. .20167)  
/note="MER33 repeat: matches 302. .26 of consensus"  
repeat\_region  
24403. .24428  
/note="13 copies 2 mer tt 92% conserved"  
repeat\_region  
24433. .24855  
/note="MLT2\_internal repeat: matches 1739. .2174 of consensus"  
repeat\_region  
complement(24871. .25777)  
/note="L1PA15 repeat: matches 904. .1 of consensus"  
repeat\_region  
complement(25626. .26986)  
/note="L1 repeat: matches 5390. .3989 of consensus"  
repeat\_region  
26985. .27537  
/note="L1 repeat: matches 2116. .2707 of consensus"  
repeat\_region  
complement(27553. .27820)  
/note="AluY repeat: matches 301. .34 of consensus"

repeat\_region  
2722. .29084  
/note="L1 repeat: matches 2700. .3980 of consensus"  
repeat\_region  
29060. .29679  
/note="MLT2\_internal repeat: matches 2145. .2774 of consensus"  
repeat\_region  
29763. .30148  
/note="MSTD repeat: matches 3. .394 of consensus"  
repeat\_region  
30148. .30776  
/note="MLT2\_internal repeat: matches 2859. .3495 of consensus"  
repeat\_region  
30922. .31355  
/note="MLT2\_internal repeat: matches 3710. .4142 of consensus"  
repeat\_region  
31659. .32910  
/note="L1 repeat: matches 4149. .5390 of consensus"  
repeat\_region  
32755. .33606  
/note="L1MB6 repeat: matches 1. .912 of consensus"  
repeat\_region  
complement(34384. .34589)  
/note="MIR repeat: matches 247. .32 of consensus"  
repeat\_region  
35440. .35544  
/note="MER5A repeat: matches 4. .109 of consensus"  
repeat\_region  
36090. .36149  
/note="15 copies 4 mer gaag 100% conserved"  
repeat\_region  
complement(36456. .36554)  
/note="MIR repeat: matches 146. .48 of consensus"  
misc\_feature  
37067. .37716  
/note="match: GSS B65235"  
repeat\_region  
complement(37373. .37670)  
/note="AluX repeat: matches 297. .2 of consensus"  
repeat\_region  
complement(37694. .37994)  
/note="AluSg repeat: matches 300. .1 of consensus"  
prim\_transcript  
38661. .40016  
/note="match: multiple ESTs: match: EST Z38206 clone c-02h10; Paired with EST Z41912 matching this clone; match: EST Z41912 clone c-02h10; Paired with EST Z38206 matching this clone"  
misc\_feature  
38661. .38857  
/note="match: STS G05514"  
prim\_transcript  
40911. .>42821  
/note="match: multiple ESTs: match: N50346 R49495 H21114 AA350242 H46781; match: H29818 F04199 H18678 H50505 F07947; match: H05281 F02238 N51569 H46205 H40284; match: H46206 AA968020 R53041 R16141 H21157; match: Z42769 R53128 H29905 H51336 AA332574; match: R35107 AA968020 F07947 H18769 AA350243"  
misc\_feature  
40991. .>41471  
/note="match: STS G24255"  
repeat\_region  
41587. .41620  
/note="17 copies 2 mer tg 85% conserved"  
prim\_transcript  
complement(44566. .47136)  
/note="match: 5' EST R15400 clone 53057"  
repeat\_region  
complement(46334. .46442)  
/note="MIR repeat: matches 157. .48 of consensus"  
repeat\_region  
complement(46542. .46643)  
/note="MIR2 repeat: matches 145. .41 of consensus"  
prim\_transcript  
47141. .47466  
/note="match: 3' EST AA985139 clone IMAGE:1616277"  
repeat\_region  
complement(50606. .50641)  
/note="MIR2 repeat: matches 143. .108 of consensus"  
repeat\_region  
50807. .50922  
/note="MIR2 repeat: matches 9. .128 of consensus"  
repeat\_region  
50878. .50920  
/note="MIR repeat: matches 206. .248 of consensus"  
repeat\_region  
complement(51705. .51834)  
/note="MIR repeat: matches 192. .52 of consensus"  
repeat\_region  
complement(52051. .52120)  
/note="MIR repeat: matches 150. .81 of consensus"  
repeat\_region  
complement(56281. .56405)  
/note="MIR2 repeat: matches 146. .18 of consensus"  
repeat\_region  
complement(56554. .56855)  
/note="AluY repeat: matches 301. .1 of consensus"  
repeat\_region  
complement(57195. .57494)  
/note="AluY repeat: matches 300. .1 of consensus"

	repeat_region	complement(57507..57630)	
	repeat_region	/note="MIR repeat: matches 187..66 of consensus"	
	repeat_region	complement(57973..59073)	
	repeat_region	/note="L1 repeat: matches 5047..3929 of consensus"	
	repeat_region	complement(60809..61101)	
	repeat_region	/note="AlusC repeat: matches 293..1 of consensus"	
	repeat_region	complement(62404..64830)	
	repeat_region	/note="TIGER1 repeat: matches 2418..1 of consensus"	
	repeat_region	64848..64979	
	repeat_region	/note="MIR2 repeat: matches 15..146 of consensus"	
	repeat_region	65556..65790	
	misc_feature	<65829..66276	
	repeat_region	/note="match: GSS AQQ02124"	
	repeat_region	67773..67839	
	repeat_region	/note="MIR repeat: matches 79..146 of consensus"	
	repeat_region	68924..68987	
	repeat_region	/note="32 copies 2 mer gt 83% conserved"	
	repeat_region	69506..70033	
	repeat_region	/note="L1 repeat: matches 4843..5389 of consensus"	
	repeat_region	69893..70891	
	repeat_region	/note="LINA6 repeat: matches 1..1047 of consensus"	
	Query Match	48.3%; Score 362.4; DB 9; Length 100418;	
	Best Local Similarity	75.8%; Pred. No. 1.6e-65;	
	Matches 505; Conservative	0; Mismatches 146; Indels 15; Gaps 4;	
QY	52	CGACCTTCATGATAAACTTCAGTCGGATGAGAGCTGCCTCTTTATG-ATGAACAAGAAGG 110	
DB	63058	CAATCTATTGATAAATGTGAATAAATGACGAGTGCTCTTTATGAATGAGCAAGAAAG 62999	
QY	111	TGTTTTCTTGAATGGAATCTACTCCCTGGTGAAGATGCTGTGAACATTTGTGAATTCACA 170	
DB	62998	TGGTTCTTCGAGAGGAATCTACTCTGGTGATGATCTATGAATGTTGTTGAATTCACA 62939	
QY	171	AGAAAGAAATTCACAGTGTTCATATAGAGTTAGTTGATGAAGCAGTAGCAGGATTCGAGAG 230	
DB	62938	ACAAGAGTTTAGAATATTTTCATAAACTTAGTTGATAAGCAGTGCAGGATTCAGAGG 62879	
QY	231	ATCGATTCCAATTTCAAATTAAGTTCCTCTGGGTAAAATGCTATCAATGCCGTGCGCA 290	
DB	62878	ATTGACTCCAATTTTGAAGAAGTTCATTCTGTGTAATGCTATCAATAGCTGCACAC 62819	
QY	291	TGCTACAGAGAAATCATCATGAAGGAAGAGTCAATTTGATGCGCAACTTCATTGTTG 350	
DB	62818	TGCTACAGAGAAATCTCTCATGAAGGAAGAGTCCATTGATGACAGAACTTCATTGTTG 62759	
QY	351	TGCTATTTTAAGAAATTTGTCAGGACCACCCCACTTCAACAAACCATGACCTGATCAGT 410	
DB	62758	TCTTAGTTTAAGGAATTCACAGCACCCACCCCACTTCAGCAACCCACCCCTGATCAGT 62699	
QY	411	CAGGAGCCATCCACATTGAGCGNAGACCTCCAGCAGTAAAAAGATT---ATGATTTCTCT 467	
DB	62698	CAGCAGCCATCAACATCAAGCGAGGACACTCCACAGCAAAAAGATTCCAACAACTGACT 62639	
QY	468	AAAGGATCAGATGAACATTAGCATTTTTTTAAAGCAATAAAGTATTTTACGTAAGATATG 527	
DB	62638	GATGCTCAGATGATCATTAGCATTTTCTAGCATAACATATATTTAATTAAAGGTATGT 62579	
QY	528	TAAGTATTTTTTAGCATTAATGCTATTATGTCATTTTAATAGACTCCAGTATATTGTAAC 587	
DB	62578	ACATTGTTTGTTCAGACATAATGCTATTGAACACTTAATAGACTACAGTATAGTGTAAAC 62519	
QY	588	ATAACTTT--AAATGCACATGGAGATAAAGTAT-----TTGCTCTTTTATGATAT 636	
DB	62518	ATAACTTTCAATGGCAGTAGGAACCAAAAATTTGTCATGACTTTGCTTTTATTTGATAT 62459	
QY	637	TTCCTTTATTGTCAGTACTGTGAATGGAAACTACATTTATCTCTGGGTACACCTGTATAC 696	
DB	62458	TTCCTTTATTACAGTGGTCTAARAGCTGAACCCATAATATCTCTGAGGTACACCTGTATAA 62399	
QY	697	AGAAAG 702	

Db 62398 TGCCAG 62393

RESULTS

```

AC069000      158149 bp      DNA      linear      HTG 12-AUG-2000
Locus      Homo sapiens chromosome 1 clone RP11-114G14, WORKING DRAFT
Definition      SEQUENCE, 15 unordered pieces.
Accession      AC069000
Version      AC069000.2 GI:8468957
Keywords      HTG; HTGS-PHASE1; HTGS_DRAFT.
Source      Homo sapiens.
Organism      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 158149)
Waterston,R.H.
Title      The sequence of Homo sapiens clone
Journal      Unpublished
Reference      2 (bases 1 to 158149)
Authors      Waterston,R.H.
Title      Direct Submission
Journal      Submitted (16-MAY-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
Comment      On Jun 12, 2000 this sequence version replaced gi:7839915.

```

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0114C14
----- Summary Statistics -----
Sequencing vector: M13, 100%
Chemistry: Dye-terminator; plasmid; 0%
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 151997 bases at least Q40
Consensus quality: 153927 bases at least Q30
Consensus quality: 155094 bases at least Q20
Insert size: 126000; agarose-ftp
Insert size: 156749; sum-of-contigs
Quality coverage: 5.73 in Q20 bases; agarose-ftp
Quality coverage: 4.63 in Q20 bases; sum-of-contigs

```

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 15 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

	1930: contig	of 1930 bp	in length
1931	gap of	unknown	length
2031	3803: contig	of 1773 bp	in length
2331	3803: contig	of unknown	length
3804	3903: gap of	unknown	length
3904	5566: contig	of 1663 bp	in length
5567	5666: gap of	unknown	length
5667	10320: contig	of 4654 bp	in length
10321	10420: gap of	unknown	length
10421	16200: contig	of 5780 bp	in length
16201	16300: gap of	unknown	length
16301	23731: contig	of 7491 bp	in length
23732	2381: gap of	unknown	length
23892	32610: contig	of 8719 bp	in length
32611	32710: gap of	unknown	length
32711	43891: contig	of 11181 bp	in length
43892	43991: gap of	unknown	length
43992	53368: contig	of 9377 bp	in length
53369	53468: gap of	unknown	length

\* 53469 66543: contig of 13075 bp in length  
\* 66544 66643: gap of unknown length  
\* 66644 80033: contig of 13390 bp in length  
\* 80033 80133: gap of unknown length  
\* 80133 94139: contig of 14006 bp in length  
\* 94139 94239: gap of unknown length  
\* 94239 107329: contig of 13090 bp in length  
\* 107329 107430: gap of unknown length  
\* 107430 131992: contig of 24563 bp in length  
\* 131992 132093: gap of unknown length  
\* 132093 158149: contig of 26057 bp in length.

FEATURES

Source  
1..158149  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/clone="RP11-114G14"  
1..1930  
/note="assembly\_name:Contig2"  
2031..3803  
/note="assembly\_name:Contig3"  
3904..5566  
/note="assembly\_name:Contig4"  
5667..10320  
/note="assembly\_name:Contig5"  
10421..16200  
/note="assembly\_name:Contig6"  
16301..23791  
/note="assembly\_name:Contig7"  
23892..32610  
/note="assembly\_name:Contig8"  
32711..43891  
/note="assembly\_name:Contig9"  
43992..53368  
/note="assembly\_name:Contig10"  
53469..66543  
/note="assembly\_name:Contig11  
clone\_end:T7  
vector\_side:left"  
66644..80033  
/note="assembly\_name:Contig12"  
80134..94139  
/note="assembly\_name:Contig13"  
94240..107329  
/note="assembly\_name:Contig14"  
107430..131992  
/note="assembly\_name:Contig15"  
132093..158149  
/note="assembly\_name:Contig16"  
BASE COUNT 47784 a 31862 c 32435 g 44661 t 1407 others  
ORIGIN

Query Match 48.3%; Score 362.4; DB 2; Length 158149;  
Best Local Similarity 75.8%; Pred. No. 1.5e-65;  
Matches 505; Conservative 0; Mismatches 146; Indels 15; Gaps 4;  
QY 52 CGACATTCATGATAAATCTCAGTGGATGAGGAGCTGCTCTTATG-ATGACAAGAAGG 110  
Db 136204 CAATCTATGATAAATGTAATAATGACGAGTGTCTTATGATGACGAAGAAG 136263  
QY 111 TGGTTCTTGAATGAATCTACTCTCTGTTGAAGATCTGTGAACATTGTTGAATGACA 170  
Db 136264 TGGTTCTTGAAGGAATCTACTCTGTTGATGATGATGATGATGTTGTTGAATGACA 136323  
QY 171 AGAAGAATTTACAGTTTACATAGAGTTAGTTGATGAAGCAGTAGCAGGATTCGAGAGG 230  
Db 136324 ACAAAGAGTTTGAATATTCATAAACTTAGTTGATAAGCAGTGGCAGGATTCAGAGG 136383  
QY 231 ATCGATTCAATTTCAAAATAGTTCTCTGTTGGTAAATGCTATCAAAATGGCGTCCGA 290  
Db 136384 AFTGACTCAATTTTGAAGAAGCTTCTATTTGTTGGTAAATGCTATCAAAATAGGCTACA 136443  
QY 291 TGCTACAGAGAATCTATCATGATGAAGAAGCAGTCAATTGATGTGGCAACATTCATTGTTG 350

Db 136444 TGCTACAGAGAATCTCATGAAAGGAAGAGTCCATTGATGAGCAAACTTCATTGTTG 136503  
QY 351 TCGTATTTTAAAGAAATTTGTCAGGACCAACCCCAACCTTCAACAACCATGACCTGATCAGT 410  
Db 136504 TCTTAGTTTAAAGAAATTTGCCACAGCCACCCCAACCTTTCAGCAACCAACCCCTGATCAGT 136563  
QY 411 CAGGAGCCATCCACATTCAGGCGGAGAACCTCCAGCAGTAAAAAGATT---ATGATTCTCT 467  
Db 136564 CAGCAGCCATCAACATCAAGGCGAGGACACTCCACCACGAAAAAGATTCCAACAACCTGACT 136623  
QY 468 AAAGATCAGATGAACATTAAGCATTTTAAAGCAATAAAGTATTTTACGTAAGATATG 527  
Db 136624 GATGGCTCAGATGATCATTAAGCATTTTCTAGCAATAAATATTTATTAATTAAGGATGT 136683  
QY 528 TATGTTATTTTATAGGCAATATGCTATTTATGATTTTAAAGATCCTCAGTATATTTGTAAC 587  
Db 136684 ACATTGTTTGTTCAGACATAATGCTATTGACACATTAATAGACTACAGTACTAGTAAAC 136743  
QY 588 ATAACCTTT--AAATGCACTGGGAGATATAAGTAT-----TTGCTCTTTTATGATAT 636  
Db 136744 ATAATTTTCATATGCACTAGGAAACCAAAAATATGTCATGACTTGCCTTTATTTGATAT 136803  
QY 637 TTGCTTATTCAGTAGTCTGTAATGGAACATACATTTATCTCTGGGTACACCTGTATAC 696  
Db 136804 TTGCTTATTCAGTGGTCTTAAAGCTGAACCATATATCTCTGAGGTACACCTGTATAA 136863  
QY 697 AGAAG 702  
Db 136864 TGCCAG 136869

RESULT 10  
AC123789  
LOCUS  
DEFINITION  
Homo sapiens chromosome 11 clone RP5-998N23 map 11, \*\*\* SEQUENCING  
IN PROGRESS \*\*\*, 3 ordered pieces.  
AC123789  
VERSION AC123789.4 GI:22297442  
KEYWORDS HTG; HTGS\_PHASE2; HTGS\_FULLTOP; HTGS\_ACTIVEFIN.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 138364)  
AUTHORS Birren,B., Nusbaum,C. and Lander,E.  
TITLE Homo sapiens chromosome 11, clone RP5-998N23  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 138364)  
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,  
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,  
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,  
Cook,A., Cooke,P., Dearrellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., Fitzgerald,M., FitzHugh,W., Gage,D.,  
Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,  
Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,  
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,  
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,  
Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N.,  
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Melidrim,J.,  
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,T., Nguyen,C.,  
Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,  
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,  
Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,  
Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,  
Schubakov,R., Seaman,S., Severy,P., Spencer,B., Stange-Thoman,N.,  
Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,  
Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,  
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (01-JUN-2002) Whitehead Institute/MIT Center for Genome

REFERENCE  
AUTHORS  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 138364)  
Birren,B., Nustbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,  
Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,  
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collimore,A.,  
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,  
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,  
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,  
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,  
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,  
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,  
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,  
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,  
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,  
Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,  
Zembek,L., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (19-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Aug 19, 2002 this sequence version replaced gi:21699405.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L27325  
Center clone name: 998\_N\_23  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 3 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1 16246: contig of 16246 bp in length  
\* 16247 16346: gap of 100 bp  
\* 16347 135511: contig of 119165 bp in length  
\* 135512 135611: gap of 100 bp  
\* 135612 138364: contig of 2753 bp in length.  
FEATURES  
source  
Location/Qualifiers  
1. .138364  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11"  
/clone="Rp5-998N23"  
/clone\_lib="RP5-998N23"  
BASE COUNT 34454 a 34988 c 41938 g 26489 t 495 others  
ORIGIN  
Query Match 48.0%; Score 360.8; DB 2; Length 138364;  
Best Local Similarity 73.3%; Pred. No. 3.3e-65;  
Matches 504; Conservative 0; Mismatches 177; Indels 7; Gaps 3;  
QY 26 GGATTACAGGATGAGCCACCATGGTCAGCTTCATGATAAAACCTTCAGTGATGAGGAC 85  
DB 15018 GGCTTAAGACATGATGAATCGCTGAGCCTCATGATAAAATTTGATCAGATGAGGAGT 15077  
QY 86 TGCTCTTTATGATGAACAAAGAGGTGTTCTTCTTGAATGAATGGAATCTACTCCGGT---GA 142  
DB 15078 TGTTCTAGGATGAGCAAGATTGCTTCTTCTGAGATGGAATCTACTCTGCTAAAGA 15137

QY 143 AGATCCTGTGAACATTGTTGAAATGACAAGAAAGAAATTTACAGTGTTCATAGAGTTAGT 202  
DB 15138 AGATCAGTGAATATTGTTGAAATGACAACAAGGGATTGGAATATGCGCTCAACTTAGT 15197  
QY 203 TGATGAACAGTACAGGAGGATTCGAGAGGATCGATTCCAATTTCAAAATAAGTTCTTCTGT 262  
DB 15198 TGATAAGCAGTGGAGGGTTTGAGAGGCTTGATTCCAATTTTGAAGAAATTTCTACTGT 15257  
QY 263 GGGTAAATGCTATCAAAATGGGCGTCCTACAGAGAATCTATCATGATGAAGGAAGAG 322  
DB 15258 GGTAAATGCTATCAACAGCATCATGCTACAGAGAATCTTCCATAAAGGAATAG 15317  
QY 323 TCAATTGATGTGGCAAACTTCATTGTTGCTGTTTAAAGAAATTTGTCAGGACCAACCCCA 382  
DB 15318 CCGATCAATGTGGCAAACTTCCTGTTGCTTATTTAAGAAATTTGCCACAGCAACCCCA 15377  
QY 383 ACCTTCAACAAACCATGACCCCTGATCAGTCAGGAGCCATCCACATTTGAGGCGAGAACCTCC 442  
DB 15378 TCCTTCAGCAACCAACCCCTGATCAGCAGCAGCCATCAATACTAGGCAAGACCCCTCC 15437  
QY 443 AGCAGTAAAGATATGATTCTCTAAGAGATCAGATGAACATTTAGCATTTTAAAGCA 502  
DB 15438 ACCAGCAAAAGATCATGACTCCTTGAAGGCTCAGATGATCATTAGCA-TTTTTAGCGA 15496  
QY 503 ATAAAGTATTTTACGTAAGATGATGTTATTTTGTAGCATAATGCTATTATGCTATT 562  
DB 15497 TACATTTTAACTAGAGTCTATATATGTTTGTAGCACATAATCTATTGACACAT 15556  
QY 563 TAATAGATCCAGTATATTGTAACATAACTTTAAATGCACTGGGAGATAAAGT---AT 619  
DB 15557 TAATAGCTGTAGTATAGTATAAGCATATATGCACTGGGAAGCAAAAAAATGTTTAC 15616  
QY 620 TTGCTCTTTATGATATTTGCTTTTATGCACTAGTCTGTTATGGAAGAACTACATATCTCT 679  
DB 15617 TTGCTTTATGCAATATTTGCTTTATTTGTTGCTCTGCTCAACTGAACCTGCAATAGCTTT 15676  
QY 680 TGGGTACACCTGCTATACAGAAAGAAAT 707  
DB 15677 GAGGCATTCCTGTAGAGTGGAGAATGTT 15704  
RESULT 11  
AC124056  
LOCUS  
DEFINITION Homo sapiens chromosome 11 clone RP5-1082L12 map 11, \*\*\* SEQUENCING  
IN PROGRESS \*\*\*, 2 ordered pieces.  
AC124056  
AC124056 4 GI:22213457  
VERSION  
KEYWORDS HTG; HTGS\_PHASE2; HTGS\_FULLTOP; HTGS\_ACTIVEFIN.  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 115644)  
Birren,B., Nustbaum,C. and Lander,E.  
Homo sapiens chromosome 11, clone RP5-1082L12  
Unpublished  
2 (bases 1 to 115644)  
Birren,B., Linton,L., Nustbaum,C., Lander,E., Ali,A., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,  
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,  
Chazaro,B., Choepel,Y., Collimore,A., Collins,S., Collimore,A.,  
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,  
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,  
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,  
Karatas,A., Kells,C., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,  
Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N.,  
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,  
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,  
Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,  
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,  
Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A.,  
Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B.,  
Stange-Thomann,N., Stojanovic,N., Talamas,J., Testaye,S., Theodore,J.,  
Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,  
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (09-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 115644)

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgaiter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (13-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Aug 13, 2002 this sequence version replaced gi:21699383.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L27400  
Center clone name: 1082\_L\_12

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* been provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1 24834: contig of 24834 bp in length  
\* 24835 24934: gap of 100 bp  
\* 24935 115644: contig of 90710 bp in length.

Location/Qualifiers  
1. 115644  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11"  
/clone="RP5-1082L12"  
/clone\_lib="RPDI Human PAC library 5"  
28356 a 29538 c 29243 g 28407 t 100 others

EASE COUNT 28356 a 29538 c 29243 g 28407 t 100 others  
ORIGIN  
Query Match 47.2%; Score 354.2; DB 2; Length 115644;  
Best Local Similarity 76.7%; Pred. No. 7.8e-64;  
Matches 501; Conservative 0; Mismatches 138; Indels 14; Gaps 5;  
QY 52 CGACTTCATGATAAACTTCAGTGGATGAGGAGCTGCCTCTTAT-GATGAACAAGAGG 110

Db 30290 CAATCTCATGATCAAACTTTGAACAGATGAGGAGTTGCTTCTTTATGGATAAGCAAGAAAG 30349  
QY 111 TGGTTTCTTGAATGGAATCTACTCTCTGTGTGAAGATGCTGTGAACATTTGTGAATGACA 170  
Db 30350 TGGTTTCTTGAATGGAATCTACTCTGTGTGAAGATGCTGTGAACATTTGTGAATGACA 30409  
QY 171 AGAAGAATTTACAGTGTTTACATAGAGTTAGTTGATGAAGCAGTAGCAGGATTCGAGAGG 230  
Db 30410 ACAAAGGATTTAGAATATTCCTAATAACACAGCTGATAAAGCAGCAGCAGGTTTGAGAGG 30469  
QY 231 ATCGATTCCTCAATTTCAAAATAGTTCTTCTGTGGTGAATGCTATCAAAATGGCGTCGCA 290  
Db 30470 ATTGACTTCAGTTTGTGAAGAAGTTCTACTGTGGATGGAATGCTGTCAAAACAGCATCACA 30529  
QY 291 TGCTACAGAGAAATCTATCATCAAGAGAGAGTCAATTCATGTGGCAAACTTCATTTGTTG 350  
Db 30530 TGCTACAGAGAAATCTTTTGTGAAGGAGAGTCAATTCATGTGGCAAACTTCATTTGTTG 30589  
QY 351 TCGTATTTTAAGAATTTCTCAGGACACCCCAACCTTCAACCAACCATGACCCGTGATCAGT 410  
Db 30590 TCCTACTTTACAAATTTGCTATGTCACCCCAACCTTCAGCAATCACCAGCTGATCAAT 30649  
QY 411 CAGGAGCATCCACATTTGAGGAGAGCTCCAGCAGTAAAGAGATTTATGATTTCTCTAAA 470  
Db 30650 CGGCAGCATCAACACATCAGGCAAGACCCCTCTATCAGCAAAAGAGATTTATGACTCACTGAA 30709  
QY 471 GGATCAGATGAACATTTAGCATTTTAAAGCAATTAAGTATTTTACGTGAAGATATGTAT 530  
Db 30710 GGCTCATATGATTTGTTAGCA--TTTGAAGCAATTAACGTATTTTAA--AATTAAGATAT 30764  
QY 531 GTTATTTTATGAGCATATGCTATTTATGATTTATGATTTATGATTTATGATTTATGATTTAT 590  
Db 30765 GTACATTTTATGACATATGCTATTTGCAATTAATATGATGATATGATATGATATGATAT 30824  
QY 591 AC-TTTAAATGCACCTGGGAGATAAAAGTAT-----TTGCTCTTTTATGATATTTGCTT 642  
Db 30825 ACTTTTATGCTCTGAGAAACCAAAATATGTCATGCTTTTATGATGATTTGATTTGATTTGATTT 30884  
QY 643 TATTGCACTAGTCTGTGTAATGAAACTACATTTATCTTCTTTGGGTACACCTGTATA 695  
Db 30885 TATTGCTGTGCTGTGAACCAAGCCACCAATATACCGAGGTATGCTGTATA 30937

RESULT 12  
AC004736 115958 bp DNA linear PRI 03-JUL-1998  
LOCUS Human Chromosome 11p14.3 PAC clone pDJ1082L12 containing KNCN1 and  
DEFINITION MyoD, complete sequence.  
ACCESSION AC004736  
VERSION AC004736.1 GI:3282183  
KEYWORDS HTG.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 115958)  
AUTHORS Evans, G.A., Athanasiou, M., Aguayo, P., Armstrong, D., Basit, M., Buettner, J., Bumeister, R., Card, P., deSailboat, F., Dunn, J., English, C., Ethridge, S., Garner, H.R., Gee, V., Gordon, M., Gotway, G., Grant, O., Hahner, L., Harris, J., Lewis, E., Loo, H., Loo, K.N., Major, T., McFarland, J., Newton, J., Osborn-Lawrence, S., Schageman, J., Schultz, R.A., Stimson, S., Syed, M. and Ward, T.  
TITLE HTGS Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 115958)  
AUTHORS Evans, G.A., Athanasiou, M., Aguayo, P., Armstrong, D., Basit, M., Buettner, J., Bumeister, R., Card, P., deSailboat, F., Dunn, J., English, C., Ethridge, S., Garner, H.R., Gee, V., Gordon, M., Gotway, G., Grant, O., Hahner, L., Harris, J., Lewis, E., Loo, H., Loo, K.N., McFarland, J., Miller, R., Newton, J., Osborn-Lawrence, S., Schageman, J., Schultz, R.A., Stimson, S., Syed, M. and Ward, T.  
TITLE Direct Submission

JOURNAL Submitted (23-MAY-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA

REFERENCE 3 (bases 1 to 115958)

AUTHORS Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buettner,J., Bumeister,R., Card,P., deSailboat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.

TITLE Direct Submission

JOURNAL Submitted (01-JUL-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA

REFERENCE 4 (bases 1 to 115958)

AUTHORS Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buettner,J., Bumeister,R., Card,P., deSailboat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.

TITLE Direct Submission

JOURNAL Submitted (03-JUL-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA

COMMENT On Jul 1, 1998 this sequence version replaced gi:3152625. NOTE: Extent of pDJ1082L12 overlap with mapped clones; SP6 PAC cloning end: 1..8775, pDJ59m18: T7 PAC cloning end: 45252..115958, 6-109h6. Further information regarding the map of this region or annotation of pDJ1082L12 can be found at http://gestec.swmed.edu/chromosome2.htm.

CHROMOSOMAL LOCUS: This PAC clone comes from the Usher syndrome region (USH1C) mapped between STS markers D11S1310 and D11S1414. MARKER CONFIRMATION: STSs/ESTs sequence confirmed; D11S4130, KNCN1, and Myo1.

MAPPED CLONE OVERLAP: pDJ59m18 and 6-109h6. IMPORTANT: This submission contains the entire insert of clone pDJ1082L12. pDJ1082L12 comes from a PAC library constructed at the Roswell Park Cancer Institute by the Pieter de Jong group. This clone has been finished according to strict quality criteria and attempts have been made to resolve all base calling problems such as compressions and repetitive elements. The expected Phred/Phrap calculated errors/10kb is 0.39. In addition, this sequence has been finished such that 99.9% of consensus base calls consist of either double-stranded coverage or 2 types of labeling chemistry on one strand.

FEATURES

Source	Location/Qualifiers
repeat_region	1..115958 /organism="Homo sapiens" /db_xref="taxon:9606" 4409..4517
repeat_region	/rpt_family="MIR" complement(19171..19291)
repeat_region	/rpt_family="Alu" complement(21561..21985)
repeat_region	/rpt_family="Alu" 23678..23767
repeat_region	/rpt_family="MIR" 28877..31250
repeat_region	/rpt_family="TigterI" complement(31777..31876)
repeat_region	/rpt_family="MIR" complement(36023..36231)
repeat_region	/rpt_family="MIR" complement(39712..39766)
repeat_region	/rpt_family="MIR" complement(41971..42073)
repeat_region	/rpt_family="MIR" 45642..45752
repeat_region	/rpt_family="MIR" 57204..57337
repeat_region	/rpt_family="MIR" complement(58027..58249)

repeat_region	/rpt_family="MER20" 59316..59729
repeat_region	/rpt_family="Li" complement(60013..60386)
repeat_region	/rpt_family="THE1" 60398..60596
repeat_region	/rpt_family="Li" 60940..61228
repeat_region	/rpt_family="Alu" complement(63635..63900)
repeat_region	/rpt_family="Alu" complement(75510..75774)
repeat_region	/rpt_family="Alu" complement(82106..82378)
repeat_region	/rpt_family="Alu" complement(84150..84211)
repeat_region	/rpt_family="MLT1" 84727..85172
repeat_region	/rpt_family="MLT1" 85852..86134
repeat_region	/rpt_family="Alu" 86387..86565
repeat_region	/rpt_family="Alu" complement(88437..88514)
repeat_region	/rpt_family="MIR" complement(90686..90752)
repeat_region	/rpt_family="MIR" 91042..91163
repeat_region	/rpt_family="MIR" 94274..94419
repeat_region	/rpt_family="MIR" complement(95470..95617)
repeat_region	/rpt_family="MLT1" complement(97484..97917)
repeat_region	/rpt_family="Li" 105386..105657
repeat_region	/rpt_family="MIR" 105673..105884
repeat_region	/rpt_family="MIR" complement(106369..106641)
repeat_region	/rpt_family="Alu" complement(107302..107596)
repeat_region	/rpt_family="Alu" 109188..109458
repeat_region	/rpt_family="Alu" complement(110473..110773)
repeat_region	/rpt_family="Alu" complement(111227..111337)
repeat_region	/rpt_family="Li" 111412..111792
repeat_region	/rpt_family="THE1" complement(112629..112992)
repeat_region	/rpt_family="Li"
BASE COUNT	28386 a 29554 c 29450 g 28568 t
ORIGIN	
Query Match	47.2%; Score 354.2; DB 9; Length 115958;
Best Local Similarity	76.7%; Pred No. 7.8e-64;
Matches	501; Conservative 0; Mismatches 138; Indels 14; Gaps 5;
QY	52 CGACTTCATGATAAACTTCACTGGATGAGGAGCTGCTCTTAT-GATGAACAAGAAGG 110
Db	30607 CAATCTCATGATCAAACTTGAACAGATGAGAGTTCCTCTTATGGATAACCAAGAAG 30666
QY	111 TGGTTTCTTGAATCGAATCTACTCTCTGGTGAAGTCTGTGAACATTTGTTGAATGACA 170
Db	30667 TGGTTTCTTGAATCGAATCTACTCTCTGGTGAAGTCTGTGAACATTTGTTGAATGACA 30726
QY	171 AGAAGAAATTTACAGTGTACATAGATTAGTTAGTGAAGCAGTAGCAGATTCCGAGG 230
Db	30727 ACAAGGATTTAGATATATTCATAAAGCAGCTGATTAAGCAGCAGGTTTGAGG 30786
QY	231 ATCGATTCCAATTTCAAAATAAGTTCTTCTGTGGGTAATAATGCTATCAATATGCGTCGCA 290

REFERENCE AUTHORS TITLE JOURNAL	REFERENCE AUTHORS TITLE JOURNAL	REFERENCE AUTHORS TITLE JOURNAL	COMMENT
1 1 1 1	1 1 1 1	1 1 1 1	

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: [http://qc.bmcnrc.edu:8088/quality\\_info/geobackannotation](http://qc.bmcnrc.edu:8088/quality_info/geobackannotation)

QUALSTAT-REPORT-----  
 http://www.cmc.edu/0000/quality.html.

```
----- Summary Statistics -----
Contig length:      161348
Phread values in contigs: 16929
```

```

Phrap values in estimate:      160601
Average error rate (BCM-Phrap estimate): 0.000327607
Fraction of Phrap values less than 40 : 0.038275
Number of consensus changing edits:      57
Number of N's in consensus :             0

```

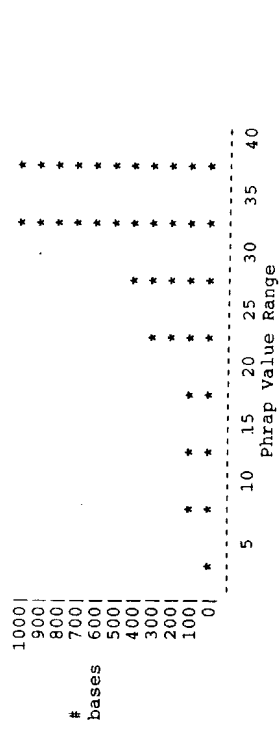
Position	Original+Context	Edited+Context
15946	ggaactgc(a)ccttaagcga	ggaactgc(a)ccttaagcga
20708	tgtgaatat(n)acagattgnc	tgtgaatat(g)acagattgac
20717	tgaagattt(n)aaacattcnc	tgaagattt(n)aaacattcnc

$\gamma_{\text{acc}}^{\text{acc}}(n)/\gamma_{\text{acc}}^{\text{acc}}(a)$



[illegible]

----- Distribution of Quality &lt; 40 Bases -----



-----					
FEATURES	Version:	1.01	qxfo.		
source	Location/Qualifiers				
	1. .161339				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/chromosome="12"				
	/clone="RPC111-154121"				
misc_feature	1. .30417				
	/note="Overlaps bases 74213 to 104630 in AC007397"				
	/function="Overlap with adjacent clone AC007397"				
repeat_region	234..372				
	/rpt_family="L2"				
repeat_region	447..521				
Query Match	47.1%	Score	354;	DB	9; Length 161339;
Best Local Similarity	75.6%;	Pred. No.	8.le-64;		
Matches	493;	Conservative	0;	Mismatches	150; Indels 9; Gaps 4;
Qy	48	TGTCGACTTTCATGATAAACTTCAGTGGATGAGGAGTGCTCTTA-TGATGAACAAGAAG	106		
Db	65322	TGCTCAAACTCATGATAAACTTTAGGAATCAGGAGTTGATTCTTACAGATGACGAAG	65381		
Qy	107	AAGGTGGTTTCTTGAATGGAATCTACTCTGTGTGAAGATGCTGTGAACATTTGTTGAAT	166		
Db	65382	AAAGTGTTTTCTTGAGATGAAATCTATTCTAATGAAGATGCCAT-AAcATTGTCGAAT	65440		
Qy	167	GACAGAAGAATTTACAGTGTTACATAGAGTTAGTTGATGAAGCAGTAGCAGGATTCGA	226		
Db	65441	AACGAAAAGAATTTAGAATATTACATAAACTTAGTTGATAAAGCAGTAGTAGGTTTGA	65500		
Qy	227	GAGGATCGATCCAAATTTCCAATAAAGTTCTCTCTGGGTAAAAATGCTATCAAAATGGCGT	286		
Db	65501	GAGGATGCCTCCAAATTTTGAAGGAAGTTCTGTTGGGTAAAAATGCTATCAAAATATGT	65560		
Qy	287	CGCATGCTACAGAGAATCTATCATGAAGGAAGAGTCAATTGATGGCAAATTCATT	346		
Db	65561	AGCATGCTACAGAGAATCTTTCATGAAGGAAGAGTCAATTGTTGGCAAATTCAGT	65620		
Qy	347	GTTGTCGTATTTAAGAAATTTGTCAGGACCACCCCACCTTCAACACCATGACCCCTGAT	406		
Db	65621	GTTGTTTTATTTTAAAGAAATGCCATAGTCACCCCATCTTTCAGCAACTACTACCCCTGAT	65680		
Qy	407	CAGTCAGAGCCATCCACATTCAGGGGAGAACCTCCAGCAGTAAAAAGATTATGATTC	466		
Db	65681	CAGTCAGAGCCATCCACGTGGGCGGAGATCCCTCCAGCAAAAAAGATCACAACTCAC	65740		
Qy	467	TAAAGGATTCAGATGAACATTAGCATTTTTTTAAGCAATAAAGTATTTTTTACGTAAGATAT	526		
Db	65741	TGGAGGCTCAGATGATCCTTACATTTTTCTAGCAAGCAGATTTTTTTAAATTAAGGTATA	65800		
Qy	527	GTATGTTATTTTTTGGGCATAATGCTATTATGCAATTAATAGACTCCAGATATATTGAAA	586		
Db	65801	TGAATTTGATTTTTAGACATAATGCTATTGTACACTTAACAGACTACAGTATAGTATAAA	65860		
Qy	587	CATAAC-TTTAAATGCATCGGAGATATAAAGTATTTG-----CTCTTTTATGATATTTT	639		
Db	65861	TATAACTTTTATATGCACTTGGAAACAAAAATTTGTGTGACTTTTTTATCATATACTTG	65920		
Qy	640	CTTTATTGCACTAGTCTGATGGAACCTACATTAATCTCTTGGGTACACCTG	691		
Db	65921	CTTTATTGTTGGTGGTCTAGAACCAAACTGTAATATCTTCTCTAGGTACACCTG	65972		

RESULT 14  
AC092165/c  
LOCUS  
AC092165  
DEFINITION  
AC092165  
ACCESSION  
AC092165.4  
VERSION  
GI:18056750  
HTG.  
KEYWORDS  
Homo sapiens.  
SOURCE  
linear  
DNA  
10020 bp  
BAC clone RP11-753F4  
from 2,  
complete sequence.  
PRI 21-FEB-2002

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 100020)  
AUTHORS Sulston, J.E. and Waterston, R.  
TITLE Toward a complete human genome sequence  
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)  
MEDLINE 99063792  
PUBMED 9847074  
REFERENCE 2 (bases 1 to 100020)  
AUTHORS Abbott, S. and Haglund, K.  
TITLE The sequence of Homo sapiens BAC clone RP11-753F4  
JOURNAL Unpublished (2001)  
REFERENCE 3 (bases 1 to 100020)  
AUTHORS Waterston, R.H.  
TITLE Direct Submission  
JOURNAL Submitted (22-JUN-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
REFERENCE 4 (bases 1 to 100020)  
AUTHORS Waterston, R.H.  
TITLE Direct Submission  
JOURNAL Submitted (04-JAN-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
REFERENCE 5 (bases 1 to 100020)  
AUTHORS Waterston, R.H.  
TITLE Direct Submission  
JOURNAL Submitted (21-FEB-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
COMMENT On Jan 4, 2002 this sequence version replaced gi:17981677.  
----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc  
Contact: sapiens@wustl.wustl.edu  
----- Summary Statistics  
-----  
Center project name: H\_NH0753F04  
-----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:  
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J. J., and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at http://www.chori.org  
VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:  
The clone sequenced to the left is RP11-762N20, 2000 bp overlap; the clone sequenced to the right is RP11-485D7. Actual start of this clone is at base position 98039 of RP11-762N20; actual end is

at base position 100020 of RP11-753F4.  
Data from AC09338 and AC074188 was used to finish this clone, AC092165. Polymorphisms have been identified between AC074188 and AC092165.  
FEATURES  
source Location/Qualifiers  
1..100020  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="2"  
/map="2"  
/clone="RP11-753F4"  
/clone\_lib="RPCI-11"  
31..176  
/rpt\_family="L1"  
177..459  
/rpt\_family="Alu"  
/rpt\_family="Alu"  
456..501  
/rpt\_family="(CA)n"  
501..3543  
/rpt\_family="L1"  
4004..4114  
/rpt\_family="CR1"  
4631..4933  
/rpt\_family="Alu"  
6357..6664  
/rpt\_family="Alu"  
6459..6464  
/note="similar to Homo sapiens EST B1830777 (NID:gl15942327)"  
misc\_feature 8893..9252  
/note="similar to Sus scrofa EST B1346780 (NID:gl15040069)"  
misc\_feature 8894..9319  
/note="match to EST A1625488 (NID:g4650419) ty56a01.x1"  
misc\_feature 8894..9258  
/note="match to EST AW247923 (NID:g6591011)"  
misc\_feature 8896..9178  
/note="match to EST BE782599 (NID:gl0203797)"  
misc\_feature 8897..9319  
/note="similar to Rattus norvegicus EST AW523743 (NID:gl166128)"  
misc\_feature 8907..9319  
/note="similar to Mus musculus EST BB187253 (NID:gl6270533)"  
misc\_feature 8908..9319  
/note="match to EST BF663245 (NID:gl11937127)"  
misc\_feature 8920..9319  
/note="match to EST AA523527 (NID:g2264239) ni45g12.s1"  
misc\_feature 9187..9319  
/note="similar to Rattus norvegicus EST BF562978 (NID:gl1672708)"  
misc\_feature 9274..9319  
/note="similar to EST BG803768 (NID:gl17950677)"  
misc\_feature 9323  
/note="match to EST AW247923 (NID:g6591011)"  
misc\_feature 9466..9547  
/note="match to EST AA523527 (NID:g2264239) ni45g12.s1"  
misc\_feature 9466..9545  
/note="similar to Mus musculus EST BB187253 (NID:gl6270533)"  
misc\_feature 9466..9542  
/note="match to EST A1625488 (NID:g4650419) ty56a01.x1"  
misc\_feature 9466..9542  
/note="match to EST BF663245 (NID:gl11937127)"  
misc\_feature 9466..9542  
/note="match to EST BF663245 (NID:gl11937127)"  
misc\_feature 9466..9542  
/note="similar to EST BG803768 (NID:gl17950677)"  
misc\_feature 9466..9542  
/note="similar to Rattus norvegicus EST BF562978 (NID:gl1672708)"  
misc\_feature 9466..9523  
/note="similar to Rattus norvegicus EST AW523743 (NID:gl166128)"  
misc\_feature 9785..9880





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 14:50:38 ; Search time 124.796 Seconds  
(without alignments)  
13552.085 Million cell updates/sec

Title: US-09-914-152-3\_COPY\_2000\_2750  
Perfect score: 751  
Sequence: 1 cctgctgggtcccaag.....tgatggaggctggaagtcc 751

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	751	100.0	10562	21 AAA93876	Human beta3Gal-T5
2	350.6	46.7	32192	22 AAL04361	Human reproductive
3	341.8	45.5	32190	22 AAK90524	Human digestive sy
4	330.8	44.0	14117	24 ABL64107	Breast cancer rela
5	322	42.9	18765	24 AAD26402	Human glutamate re
6	315.4	42.0	112190	22 AAH44801	Human GPCR protein
7	314	41.8	236303	22 AAS11614	Human genomic DNA
8	309.4	41.2	140036	24 AAS98600	Human genomic DNA
9	305.8	40.7	143306	24 ABK49586	Human transporter

10	302.2	40.2	6022	22	ABA18218	Human nervous syst
11	301	40.1	460	24	ABL65786	Lung cancer relate
12	300	39.9	4503	22	AAL05284	Human reproductive
13	294.6	39.2	143068	21	AAF21105	Human low adenosi
14	294.6	39.2	143068	21	AAF21105	Human low adenosi
15	294.6	39.2	143068	21	AAF21105	Human low adenosi
16	294.6	39.2	143068	21	AAF21105	Human low adenosi
17	294.6	39.2	143068	21	AAF21105	Human low adenosi
18	294.6	39.2	143068	21	AAF21105	Human low adenosi
19	294.6	39.2	143068	21	AAF21105	Human low adenosi
20	293.2	39.0	160771	24	ABQ88179	Human osteoblast d
21	289.8	38.6	512	23	AAS76650	DNA encoding novel
22	287.6	38.3	72215	22	ABQ86832	Human immune/haema
23	285.8	38.1	83945	24	ABQ88101	Human immune/haema
24	280.8	37.4	147309	24	ABK49450	Human transporter
25	280	37.3	15809	22	AAK68321	Human transporter
26	278.8	37.1	814	22	AAH04355	Human cDNA clone (
27	278.8	37.1	2460	22	AAH17279	Human cDNA sequenc
28	275.2	36.6	4320	22	AAK90584	Human digestive sy
29	275.2	36.6	4320	22	AAK90585	Human digestive sy
30	274.2	36.5	125439	24	ABQ88177	Human bone marrow
31	273.2	36.4	489	22	AAK30185	Human nervous syst
32	272	36.2	32249	22	ABA07406	Human pancreatic c
33	272	36.2	32249	22	ABA07406	Human pancreatic c
34	272	36.2	32249	22	AAK91137	Human digestive sy
35	272	36.2	50442	22	AAK73083	Human immune/haema
36	272	36.2	50442	22	AAK7551	Human immune/haema
37	270	36.0	705	24	ABL87872	Human ovarian canc
38	261.6	34.8	27976	22	AAK70187	Human immune/haema
39	261.6	34.8	27976	22	AAK79311	Human immune/haema
40	256.4	34.1	1405	24	ABA05529	Molecular chaperon
41	256	34.1	109201	24	ABQ88125	Human osteoblast d
42	253	33.7	32249	22	AAL04789	Human reproductive
43	253	33.7	32249	23	ABL97684	Human testicular a
44	251.8	33.5	37808	20	AAK02780	Vector pMX-BG DNA
45	251.8	33.5	48727	22	AAK67375	Human immune/haema

ALIGNMENTS

RESULT 1  
AAA93876  
ID AAA93876 standard; DNA; 10562 BP.  
AC AAA93876;  
XX  
XX 15-JAN-2001 (first entry)  
XX Human beta3Gal-T5 encoding DNA.  
DE Human beta3Gal-T5 encoding DNA.  
KW Beta-1.3 galactose transferase; treatment; diagnosis; cancer; human;  
KW digestive system; beta3Gal-T5; ds.  
XX  
OS Homo sapiens.  
XX  
XX WO2000050608-A1.  
XX  
PD 31-AUG-2000.  
XX  
PF 24-FEB-2000; 2000WO-JP01070.  
XX  
PR 25-FEB-1999; 99JP-0047571.  
XX  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
XX Narimatsu H, Isshiki S, Togayachi A, Sasaki K;  
WPI; 2000-549409/50.  
XX Beta-1.3 galactose transferase and DNA encoding it, useful for  
PT synthesis of type 1 sialyl Lewis, a carbohydrate for treatment of  
PT digestive system cancer

XX PS Claim 31; Page 103-111; 123pp; Japanese.

XX CC This invention relates to a polypeptide (I) with beta-1.3 galactose

CC transferase activity, or variants of (I) comprising amino acid additions,

CC deletions and/or substitutions. Included in the invention is DNA encoding

CC all or part of (I); expression vectors containing the DNA, host cells

CC transformed by the vectors; a method for the preparation of the

CC polypeptide by culture of the transformants or by expression in the milk

CC of a transgenic mammal, and antibodies recognising (I). The Beta-1.3

CC galactose transferase protein transfers galactose by beta-1.3 bonding to

CC N-acetylglucosamine present in a non-cyclic carbohydrate chain (such as

CC GlcNAc) to give Galbeta1-3GlcNAc. The protein and

CC DNA encoding it are useful for the treatment and diagnosis of cancer of

CC the digestive system. The present sequence represents a Beta3Gal-T5

XX CC encoding DNA sequence.

XX SQ Sequence 10562 BP; 2610 A; 2415 C; 2574 G; 2963 T; 0 other;

Query Match 100.0%; Score 751; DB 21; Length 10562;

Best Local Similarity 100.0%; Pred. No. 6.3e-177;

Matches 751; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCGCTGGCTCCCAAGTCTAGGATTACAGGGATGAGCCACCATGGTCGACTTCAT 60

Db 2000 CCGCTGGCTCCCAAGTCTAGGATTACAGGGATGAGCCACCATGGTCGACTTCAT 2059

Qy 61 GATAAACTTCAGTGGATGAGGAGCTCCCTCTTATGATGAACAAGAGGTGGTTCTTG 120

Db 2060 GATAAACTTCAGTGGATGAGGAGCTCCCTCTTATGATGAACAAGAGGTGGTTCTTG 2119

Qy 121 AAATGAATCTACTCTGCTGAAGATGCTGTGAACATTTGTTGAATGACAAAGAAGATT 180

Db 2120 AAATGAATCTACTCTGCTGAAGATGCTGTGAACATTTGTTGAATGACAAAGAAGATT 2179

Qy 181 TACAGTGTACATAGATTAGTGTAGGACAGTACGAGATTCGAGAGATCCATCCCA 240

Db 2180 TACAGTGTACATAGATTAGTGTAGGACAGTACGAGATTCGAGAGATCCATCCCA 2239

Qy 241 ATTTCAAATAAGTTCTTCTGTGGTAAATGCTATCAAAATGGCTGCGATGCTACAGAG 300

Db 2240 ATTTCAAATAAGTTCTTCTGTGGTAAATGCTATCAAAATGGCTGCGATGCTACAGAG 2299

Qy 301 AAATCTATCATGAAGAGAGTCAATGATGTGGCAAACTTCATTTGTCGTATTTTA 360

Db 2300 AAATCTATCATGAAGAGAGTCAATGATGTGGCAAACTTCATTTGTCGTATTTTA 2359

Qy 361 AGAATTCAGGACCCACCCACCTTCAACACCATGACCCCTGATCAGTCAGGAGCCAT 420

Db 2360 AGAATTCAGGACCCACCCACCTTCAACACCATGACCCCTGATCAGTCAGGAGCCAT 2419

Qy 421 CCACATTCAGGCGAGAACCTCCAGCAGTAAAGATTTATGATTTCTTAAAGGATCAGATG 480

Db 2420 CCACATTCAGGCGAGAACCTCCAGCAGTAAAGATTTATGATTTCTTAAAGGATCAGATG 2479

Qy 481 AACATTAGCATTTTTTAAAGCAATAAAGTATTTTACGTAAGATATGATGTTATTTTTT 540

Db 2480 AACATTAGCATTTTTTAAAGCAATAAAGTATTTTACGTAAGATATGATGTTATTTTTT 2539

Qy 541 AGGCATATGCTATTATGCAATTTATAGACTCCAGTATATTGTAACATAACTTTAAATG 600

Db 2540 AGGCATATGCTATTATGCAATTTATAGACTCCAGTATATTGTAACATAACTTTAAATG 2599

Qy 601 CACTGGGAGATAAAGTATTTCTCTTTTATGATATTTCTTTATTCAGTAGTCTGTAA 660

Db 2600 CACTGGGAGATAAAGTATTTCTCTTTTATGATATTTCTTTATTCAGTAGTCTGTAA 2659

Qy 661 TGGAACTACATTTATCTCTTGGGTACACCTGTATACAGAAAGAAATTTATCATGAGAAA 720

Db 2660 TGGAACTACATTTATCTCTTGGGTACACCTGTATACAGAAAGAAATTTATCATGAGAAA 2719

Qy 721 TGCTATGCAATGATGGAGGCTGGAAAGTCC 751

XX

Db 2720 TGCTATGCAATGATGGAGGCTGGAAAGTCC 2750

RESULT 2

AAJ04361

ID AAL04361 standard; DNA; 32192 BP.

XX AC AAL04361;

XX DT 21-NOV-2001 (first entry)

XX DE Human reproductive system related antigen DNA SEQ ID NO: 7049.

XX DE Human reproductive system related antigen; reproductive system disorder;

XX KW Human; reproductive system related antigen; cancer; gene therapy; ds.

XX OS Homo sapiens.

XX PN WO200155320-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01339.

XX PR 31-JAN-2000; 2000US-0179065.

XX PR 04-FEB-2000; 2000US-0180628.

XX PR 24-FEB-2000; 2000US-0184664.

XX PR 02-MAR-2000; 2000US-0186350.

XX PR 16-MAR-2000; 2000US-0189874.

XX PR 17-MAR-2000; 2000US-0190076.

XX PR 18-APR-2000; 2000US-0198123.

XX PR 19-MAY-2000; 2000US-0205515.

XX PR 07-JUN-2000; 2000US-0209467.

XX PR 28-JUN-2000; 2000US-0214886.

XX PR 30-JUN-2000; 2000US-0215135.

XX PR 07-JUL-2000; 2000US-0216647.

XX PR 07-JUL-2000; 2000US-0216880.

XX PR 11-JUL-2000; 2000US-0217487.

XX PR 11-JUL-2000; 2000US-0217496.

XX PR 14-JUL-2000; 2000US-0218290.

XX PR 26-JUL-2000; 2000US-0220963.

XX PR 26-JUL-2000; 2000US-0220964.

XX PR 14-AUG-2000; 2000US-0224518.

XX PR 14-AUG-2000; 2000US-0224519.

XX PR 14-AUG-2000; 2000US-0225213.

XX PR 14-AUG-2000; 2000US-0225214.

XX PR 14-AUG-2000; 2000US-0225266.

XX PR 14-AUG-2000; 2000US-0225267.

XX PR 14-AUG-2000; 2000US-0225268.

XX PR 14-AUG-2000; 2000US-0225270.

XX PR 14-AUG-2000; 2000US-0225447.

XX PR 14-AUG-2000; 2000US-0225757.

XX PR 14-AUG-2000; 2000US-0225758.

XX PR 14-AUG-2000; 2000US-0225759.

XX PR 18-AUG-2000; 2000US-0226279.

XX PR 22-AUG-2000; 2000US-0226681.

XX PR 22-AUG-2000; 2000US-0226686.

XX PR 23-AUG-2000; 2000US-0227009.

XX PR 30-AUG-2000; 2000US-0228924.

XX PR 01-SEP-2000; 2000US-0229287.

XX PR 01-SEP-2000; 2000US-0229343.

XX PR 01-SEP-2000; 2000US-0229344.

XX PR 01-SEP-2000; 2000US-0229345.

XX PR 05-SEP-2000; 2000US-0229509.

XX PR 05-SEP-2000; 2000US-0229513.

XX PR 06-SEP-2000; 2000US-0230437.

XX PR 06-SEP-2000; 2000US-0230438.

XX PR 08-SEP-2000; 2000US-0231242.

XX PR 08-SEP-2000; 2000US-0231243.

XX PR 08-SEP-2000; 2000US-0231244.

XX PR 08-SEP-2000; 2000US-0231413.

XX PR 08-SEP-2000; 2000US-0231414.



QY	591	AC	TTTAAATGCACTGGGAGA	-----TAAAGTATTTGCTCTTTATGATA	635
DB	8085	ACTTTTATATGCACTGGGAACCAAGCAATTTACGTGACACAAATTTGCTTTATTACAATA	8144		
QY	636	TTTGCTTTATTGCATAGTCTCTAATGGAACTACATTATCTCTGGGTACACCTGTAT	694		
DB	8145	TTCACTTTATTTTCCATTATCTCTGAATGGAACCCACCAATTATCTTTGAGGTGTGCTGTAT	8203		

RESULT 3	
AAK90524/c	
ID AAK90524 standard; DNA; 32190 BP.	
XX	
AC AAK90524;	
XX	
DT 05-NOV-2001 (first entry)	
XX	
DE Human digestive system antigen genomic sequence SEQ ID NO: 4100.	
XX	

Human; digestive system antigen; gene therapy; cancer; appendicitis; ulcerative colitis; infection; Hirschsprung's disease; chronic colitis  
digestive system disorder; Meckel's diverticulum; ds.

XX  
OS Homo sapiens.  
XX  
PN W020015314 - A2.

XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01324.

XX	31-JAN-2000;	2000US-0179065.
PR	04-FEB-2000;	2000US-0180628.
PR	24-FEB-2000;	2000US-0184664.
PR	02-MAR-2000;	2000US-0186350.
PR	16-MAR-2000;	2000US-0189874.
PR	17-MAR-2000;	2000US-0190076.
PR	18-APR-2000;	2000US-0198123.
PR	19-MAY-2000;	2000US-0205515.
PR	07-JUN-2000;	2000US-0209467.
PR	28-JUN-2000;	2000US-0214886.
PR	30-JUN-2000;	2000US-0215135.
PR	07-JUL-2000;	2000US-0216647.
PR	07-JUL-2000;	2000US-0216880.
PR	11-JUL-2000;	2000US-0217487.
PR	11-JUL-2000;	2000US-0217496.
PR	14-JUL-2000;	2000US-0218290.
PR	26-JUL-2000;	2000US-0220963.
PR	26-JUL-2000;	2000US-0220964.
PR	14-AUG-2000;	2000US-0224518.
PR	14-AUG-2000;	2000US-0224519.
PR	14-AUG-2000;	2000US-0225213.
PR	14-AUG-2000;	2000US-0225214.
PR	14-AUG-2000;	2000US-0225266.
PR	14-AUG-2000;	2000US-0225267.
PR	14-AUG-2000;	2000US-0225268.
PR	14-AUG-2000;	2000US-0225270.
PR	14-AUG-2000;	2000US-0225447.
PR	14-AUG-2000;	2000US-0225757.
PR	14-AUG-2000;	2000US-0225758.
PR	14-AUG-2000;	2000US-0225759.
PR	15-AUG-2000;	2000US-0226279.
PR	22-AUG-2000;	2000US-0226661.
PR	22-AUG-2000;	2000US-0226868.
PR	22-AUG-2000;	2000US-0227182.
PR	23-AUG-2000;	2000US-0227009.
PR	30-AUG-2000;	2000US-0228924.
PR	01-SEP-2000;	2000US-0229287.
PR	01-SEP-2000;	2000US-0229343.
PR	01-SEP-2000;	2000US-0229344.
PR	01-SEP-2000;	2000US-0229345.
PR	05-SEP-2000;	2000US-0229509.



PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
PI WPI: 2001-502630/55.  
XX  
PT Polynucleotides encoding digestive system antigens, useful for  
PT diagnosing, treating, preventing and/or prognosing disorders of the  
PT digestive system, particularly cancer and cancer metastases -  
XX  
XX Disclosure; SEQ ID NO 4100; 986pp; English.  
PS  
XX The present invention provides the protein and coding sequences of a  
CC number of human digestive system antigens. These can be used in the  
CC diagnosis, treatment and prevention of digestive system disorders,  
CC including cancer, Meckel's diverticulum, bacterial or parasitic  
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or  
CC ulcerative colitis. The present sequence is a genomic DNA fragment  
CC encoding a digestive system antigen of the invention.  
XX  
SQ Sequence 32190 BP; 10714 A; 6313 C; 5965 G; 9198 T; 0 other;

Query Match 45.5%; Score 341.8; DB 22; Length 32190;  
Best Local Similarity 75.7%; Pred. No. 5.6e-75;  
Matches 493; Conservative 0; Mismatches 142; Indels 16; Gaps 5;

QY 52 CGACTTCATGATAAACTTCAGTGGATGAGGAGCTGCTCTTAT-GATGAACAAAGAGG 110  
Db 29946 CAATCTCATGATAAACTTGGATGAATGAGGAGTGTCTTCTGGACAAACCAAG 29887

QY 111 TGGTTTCTTGAATGAATCTACTCTGTTGAAGATCTGTGAACATGTTGAAATGACA 170  
Db 29886 TGGTTTCTTGAAGATGAACTACTCTTGGTGAAGATCTGTGAACATGTTGAAATGACA 29827

QY 171 AGAAGAAATTTACAGTGTACATAGAGTAGTTGATGAAGCAGTACAGGATTCGAGAGG 230  
Db 29826 ACAAGAAATTCAGAAGATACATAAATTTAGTTGATAACAGCTGGCAGGTTTGAGAGG 29767

QY 231 ATCGATTCAATTTCAAATAAGTCTTCTGTGGTAAATGCTATCAAAATGGGTCGCA 290  
Db 29766 ATTAGCTCCAGTTTGAAGAAGTCTTACTGTGGTAAATGCTATCAAAACAGATCACA 29707

QY 291 TGTACAGAGAAATCAATCAAGAGAGAGCAATGATGATGGCAAACTTCATTTGTTG 350  
Db 29706 CATTACAGATAAATCTTTGTGAAAAGAAAGTCAATTTGTGACGCAAC-TCTTCATTG 29648

QY 351 TCGTATTTTAAGAAATGTGACAGACCAACCCACCTTCACACACCTGACCTGATCAGT 410  
Db 29647 TCATATTTTAAAGAGTTGCCACAGACCAACCCACCTTTGAGCAACCACTACCCAGATCAAT 29588

QY 411 CAGGAGCCATCCACATTTAGGGCGAGAACCTCCAGCAGTAAAAAGATTATGATTCCTCTAA 470  
Db 29587 CAGCAGCCATCAATATCAAGGCTAGCCCTCCACCACCAAAAGACCACATCTTCGAGAA 29528

QY 471 GGATCAGATGAACATTAGCATTTTTTTAAGCAATAAAGTATTTTTACGTAAGATATGTAT 530  
Db 29527 GGCTCAGATGAACATTAGCATTTTTTTAGCATTTAAATATTTTTGAATTGAGATATGTA- 29469

QY 531 GTTATTTTATGGCATAATGCTATTATGCAATTTAATAGACTCCAGTATATTGTAACATA 590  
Db 29468 --TATTTTCAGCCCAATGCTATTACACACTTACTAGATTACAATAATAATAATGTA- 29412

QY 591 ACTTTAAATGCACCTGGGAGATAAAG-----TATTTGCTCTTTTATGATATTGCTT 642  
Db 29411 --TTTATATGCTGGGAAACAAAGGTTTGTGTAAGTCTCTTTATTTTATTTGCTT 29354

QY 643 TATTGCAATGCTGTGAATGGAACACTACATTTATCTCTTGGGTACACCTGTA 693  
Db 29353 TATTGCAATGCTGTGAACCAAACTGCAATATCTCTTAAGGTATACCTGTA 29303

RESULT 4  
ABL64107  
ID ABL64107 standard; DNA; 14117 BP.  
XX  
AC ABL64107;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Breast cancer related gene sequence SEQ ID NO:2444.  
XX  
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US10838.  
XX  
PR 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 25-SEP-2000; 2000US-234567P.  
PR 25-SEP-2000; 2000US-234923P.  
PR 25-SEP-2000; 2000US-234924P.  
PR 25-SEP-2000; 2000US-235077P.  
PR 25-SEP-2000; 2000US-235082P.  
PR 25-SEP-2000; 2000US-235134P.  
PR 25-SEP-2000; 2000US-235280P.  
PR 26-SEP-2000; 2000US-235637P.  
PR 26-SEP-2000; 2000US-235638P.  
PR 27-SEP-2000; 2000US-235711P.  
PR 27-SEP-2000; 2000US-235720P.  
PR 27-SEP-2000; 2000US-235840P.  
PR 27-SEP-2000; 2000US-235863P.  
PR 28-SEP-2000; 2000US-236038P.  
PR 28-SEP-2000; 2000US-236032P.  
PR 28-SEP-2000; 2000US-236033P.  
PR 28-SEP-2000; 2000US-236034P.  
PR 28-SEP-2000; 2000US-236109P.  
PR 28-SEP-2000; 2000US-236111P.  
PR 29-SEP-2000; 2000US-236842P.  
PR 29-SEP-2000; 2000US-236891P.

PR 02-OCT-2000; 2000US-237172P.  
PR 02-OCT-2000; 2000US-237173P.  
PR 02-OCT-2000; 2000US-237278P.  
PR 02-OCT-2000; 2000US-237294P.  
PR 02-OCT-2000; 2000US-237295P.  
PR 02-OCT-2000; 2000US-237316P.  
PR 03-OCT-2000; 2000US-237425P.  
PR 03-OCT-2000; 2000US-237598P.  
PR 03-OCT-2000; 2000US-237604P.  
PR 03-OCT-2000; 2000US-237606P.  
PR 03-OCT-2000; 2000US-237608P.  
PR 01-NOV-2000; 2000US-244867P.  
PR 01-NOV-2000; 2000US-245084P.  
PA (AVAL-) AVALON PHARM.  
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
PI Soppet DR, Weaver Z;  
XX WPI; 2002-188264/24.  
XX  
PT Screening for anti-neoplastic agent involves exposing cells to a  
PT chemical agent to be tested for anti-neoplastic activity, and  
PT determining a change in expression of a gene of a signature gene set  
XX  
XX Claim 1; SEQ ID 2444; 44pp; English.  
XX  
CC The present invention describes a method (M1) for screening for an  
CC anti-neoplastic agent. The method involves exposing cells to a chemical  
CC agent to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
CC to ABL70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening  
CC an anti-neoplastic agent, and can be used for producing a product which  
CC is the data collected with respect to the anti-neoplastic agent as a  
CC result of M1, and the data is sufficient to convey the chemical  
CC structure and/or properties of the agent. M1 can be used in the  
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer.  
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
CC carcinoma, papillary carcinoma and Wilms' tumour.  
XX  
SQ Sequence 14117 BP; 3568 A; 3182 C; 3484 G; 3883 T; 0 other;  
Query Match 44.0%; Score 330.8; DB 24; Length 14117;  
Best Local Similarity 73.6%; Pred. No. 2.4e-72;  
Matches 493; Conservative 0; Mismatches 162; Indels 15; Gaps 5;  
QY 36 GATGAGCCACCATGGTCCGACTTCATGATAAACTTCAGTGGATGAGGAGTGCCTCTTA- 94  
DB 439 GATGTGACTGATTTGGCCATGTTATGCGAAACGACGGATGAGGAAGTCTCTTAT 498  
QY 95 TGATGACAAAGAGAGGTGGTTCTTCTGAAATCTACTCTCTGGTGAAGATGCTGTGAA 154  
DB 499 TCATGACAAAGAGAGGTGGTTCTTCTGAAATCTACTCTCTGGTGAAGATGCTGTGAA 558  
QY 155 CATTGTTGAAATGACAGAAAGATTTACAGTGTTCATACATAGAGTTAGTTGATGAAGCAGT 214  
DB 559 CATCACTGAAATGGCAACAAAGAGTTTGAATACATACATCAATTTAGTTAATAGGCAGT 618  
QY 215 AGCAGGATTCGAGAGGATCGATTCATTTCAAAATAGTTCTTCTGGTGGGTAATAGT 274  
DB 619 GGCAGGGTTTTCAGAGGAGTACCTAGTTTTCAGAGAAATTTCTACTGT---TAAATGCT 675  
QY 275 ATCAATGGGTCCTGCTACAGAGAAATCTATCATCAAGAGAGATCAATTTGATGTG 334  
DB 676 ATCAACAGATCATGCTACAGGGAATTTTCATGAAAGGAGAGTCAACTGATGCG 735  
QY 335 GCAACTTCATTTGTTGCTGATTTTAAAGAAATTTGTCAGGACACCCCACTTCAACAC 394

DB 736 GCAAACTTCACTGATGTCTCATTTTTCAGAAATTCGCCACAGCCACACTTAACCTTGCAGCAAT 795  
QY 395 CATGACCCCTGATCAGTCAGGAGCCATCCACATTTGAGCGGAGAACCTTCAGCAGTAAAAAG 454  
DB 796 CATCACCCCTGATCAGTCCTCAGCCATCAACATTTGAGCGGAGAACCTTCACACGAGAAAG 855  
QY 455 ATTATGATTTCTTAAAGGATCAGATGAACATTTAGCATTCTTTTAAAGCAATAAAGTATTTT 514  
DB 856 ATTATAACTTGTCTGAAGGCTCAGATGATCTCTAGCATTCTTTAGCAAGAA-AGCATTTTTA 914  
QY 515 TAGGTAAGATATGATGTTATTTTATTTAGGCATATGCTATATATGCAATTAATAGACTCCA 574  
DB 915 AAATTAAGTTATATACATTTGTTTATAGGCCATATGCTATGTTGATACACTTAATAGACTATA 974  
QY 575 GTATATTTGTAACATAA-CITTTAAATGCCTGGGAGATAAAAGTATTG-----CT 624  
DB 975 GTATAGTGTAAATATATGTTTACATACACAGAAACCAAAAGTTTGTGTGACTCACT 1034  
QY 625 CTTTATGATATTTGCTTTTATTCAGTAGTCTGTAAATGSAACATPACATTTATCTCTGGGT 684  
DB 1035 TTATGTTACATTTGCTTTTATTTGTTGGTCTAGAACTGAACATGCACATCTCTGAGGT 1094  
QY 685 ACACCTGTAT 594  
DB 1095 ATGCCTGTAT 1104  
RESULT 5  
AAD26402/C  
ID AAD26402 standard; DNA; 18765 BP.  
XX  
AC AAD26402;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Human glutamate receptor, metabotropic 3 (GRM3) gene fragment #3.  
XX  
KW Human; glutamate receptor metabotropic 3; GRM3; neuroprotective;  
KW neotropic; gene therapy; neurotransmitter; chromosome 7q21.1-q21.2;  
KW drug screening; neurological disorder; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200196350-A2.  
XX  
PD 20-DEC-2001.  
XX  
PF 18-JUN-2001; 2001WO-US19447.  
XX  
PR 16-JUN-2000; 2000US-212328P.  
XX  
XX (GENA-) GENAISSANCE PHARM INC.  
XX  
PI Finkel K, Koshy B, Tanguay DA;  
XX  
XX WPI; 2002-090198/12.  
XX  
PT New isolated polynucleotide, a polymorphic variant of glutamate  
PT receptor, metabotropic 3 (GRM3) gene for expressing GRM3 protein  
PT isoform to screen drugs to treat GRM3 activity-related disease -  
XX  
PS Example 1; Fig 3; 165pp; English.  
XX  
CC The invention relates to an isolated polynucleotide which is a  
CC polymorphic variant of glutamate receptor, metabotropic 3 (GRM3)  
CC isogene. GRM3 is a receptor for glutamate, the major excitatory  
CC neurotransmitter in the mammalian central nervous system. Human GRM3  
CC located on chromosome 7q21.1-q21.2 is expressed in human foetal and adult  
CC whole brain especially in the caudate nucleus and corpus collosum. GRM3  
CC DNA is useful in gene therapy and also for studying the expression and  
CC function of GRM3. GRM3 polypeptide is used for screening drugs. A  
CC recombinant non-human organism is used to study expression of GRM3 SG  
CC in vivo, for in vivo screening and testing of drugs targetted against



Db 39132 TGCTGATTTTGAAGAAATGCCACAGCCACACACACCTTTGGCAACCACCCCTGATAA 39073

QY 409 GTCAGGAGCCATCCACATGTAGGCGAGAGAACCTCCAGCAGTAAAGATATGATTTCTCTA 468

Db 39072 GTCAGCAGTCATCAACATCAAGCAAGACCTTCCACACAGTAAAGGCCAATATTTCATG 39013

QY 469 AAGCATCAGATGACATTTAGCATTTTAAAGCAATAAAGTATTTTACGTAAGATATGT 528

Db 39012 AAGCTCAGATGACTGTTA--ATCTTTTATTTTAAAGCCATTAAGTATTTTAAATTAACA 38955

QY 529 ATGTTATTTTGGCATAATGCTATTTATGCAATTTAATAGATCCAGTATATTGTAACA 588

Db 38954 TGATATCTTTTACACATA--TTATCACACATTTAAAGGCTACAGTAAATGTAACA 38898

QY 589 TAAC-TTAAATGACCTGGGAGATAAAAGT-----ATTGCTCTTTTATGATATTGCT 641

Db 38897 TAACTTTATATGATCGGGAACAAATATTTGATAGTAGCTTTTATGTAATTTGCT 38838

QY 642 TTATTGCAAGTCTGTAATGAAGCACTACATTAATCTCTGGGTACACCTGTAT 694

Db 38837 TTATTAGGCATTCTGGAATTGAACCCACAAATATCTCTGAGGTATACCTCTAT 38785

RESULT 7

AAS11614

ID AAS11614 standard; DNA; 236303 BP.

XX AC AAS11614;

XX 24-OCT-2001 (first entry)

XX Human genomic DNA containing exons 2-17 of the CRIM1 gene.

DE

XX

XX CRIM-1: Human; human chromosome 2p21-16.3; ophthalmological;

KW neuroprotective; renal; osteopathic; dental; vulnery; immunogen;

KW antibody; gene therapy; neurodegenerative disease; eye disorder;

KW cataract; bone morphogenic protein; BMG; renal disease; bone abnormality;

KW tooth abnormality; wound; ds.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT exon 33104..33277

FT /\*tag= a

FT /\*number= 2

FT intron 33278..77746

FT /\*tag= b

FT /\*number= 2

FT exon 77747..77989

FT /\*tag= c

FT /\*number= 3

FT intron 77990..79103

FT /\*tag= d

FT /\*number= 3

FT exon 79104..79224

FT /\*tag= e

FT /\*number= 4

FT intron 79225..101022

FT /\*tag= f

FT /\*number= 4

FT exon 101023..101144

FT /\*tag= g

FT /\*number= 5

FT intron 101145..113377

FT /\*tag= h

FT /\*number= 5

FT exon 113378..113560

FT /\*tag= i

FT /\*number= 6

FT intron 113561..115985

FT /\*tag= j

FT /\*number= 6

FT exon 115986..116183

FT /\*tag= k

FT /\*number= 7

FT intron 116184..135707

FT /\*tag= l

FT /\*number= 7

FT exon 135708..135836

FT /\*tag= m

FT /\*number= 8

FT intron 135837..146471

FT /\*tag= n

FT /\*number= 8

FT exon 146472..146628

FT /\*tag= o

FT /\*number= 9

FT intron 146629..148761

FT /\*tag= p

FT /\*number= 9

FT exon 148762..148883

FT /\*tag= q

FT /\*number= 10

FT intron 148884..150044

FT /\*tag= r

FT /\*number= 10

FT exon 150045..150254

FT /\*tag= s

FT /\*number= 11

FT intron 150255..153815

FT /\*tag= t

FT /\*number= 11

FT exon 153816..154031

FT /\*tag= u

FT /\*number= 12

FT intron 154032..158580

FT /\*tag= v

FT /\*number= 12

FT exon 158581..158802

FT /\*tag= w

FT /\*number= 13

FT intron 158803..173982

FT /\*tag= x

FT /\*number= 13

FT exon 173983..174177

FT /\*tag= y

FT /\*number= 14

FT intron 174178..181006

FT /\*tag= z

FT /\*number= 14

FT exon 181007..181129

FT /\*tag= aa

FT /\*number= 15

FT intron 181130..183612

FT /\*tag= ab

FT /\*number= 15

FT exon 183613..183800

FT /\*tag= ac

FT /\*number= 16

FT intron 183801..185152

FT /\*tag= ad

FT /\*number= 16

FT exon 185153..187765

FT /\*tag= ae

FT /\*number= 17

XX

PN WO200138519-A1.

XX 31-MAY-2001.

XX 24-NOV-2000; 2000WO-AU01435.

XX 26-NOV-1999; 99AU-0004348.

XX (UYQU ) UNIV QUEENSLAND.

XX

PI Little M, Yamada T, Holmes G, Georgas K, Kollie G, Wilkinson L;  
XX WPI: 2001-343951/36.  
XX  
PT Nucleic acids from human chromosome 2p21-16.3 and the encoded peptide,  
PT useful for preventing, diagnosing and treating e.g. eye disease,  
PT especially cataract formation -  
XX  
XX  
PS Claim 4; Fig 3; 16pp; English.  
XX  
CC The invention relates to nucleic acids from human chromosome 2p21-16.3  
CC and the encoded peptide (and mouse and chicken orthologues) that  
CC comprises a PGCCPLP group, an insulin-like growth factor binding protein  
CC (IGFBP)-like domain, cysteine-rich domains, an RGD (undefined) group  
CC and a transmembrane domain. The protein, e.g. CRIM1, interacts with  
CC peptides of the transforming growth factor superfamily. A composition  
CC comprising an expression construct comprising the nucleic acids of the  
CC invention or a mimetic which antagonises or mimics an activity of a CRIM1  
CC polypeptide may be used in a method for modulating the biological  
CC activity of a polypeptide of the bone morphogenic protein (BMP) family.  
CC In this way they may be used to prevent or treat an eye disease,  
CC especially cataract formation. They may also be used to treat  
CC neurodegenerative diseases, renal and kidney disease, bone and tooth  
CC abnormalities, wounds and skin damage, e.g. by use of the nucleic acid in  
CC gene therapy by using antibodies directed against CRIM1 polypeptides.  
CC The present sequence is a Human genomic DNA containing exons 2-17 of  
CC the CRIM1 gene.  
XX  
SQ Sequence 236303 BP; 70199 A; 46544 C; 47996 G; 71563 T; 1 other;  
Query Match 41.8%; Score 314; DB 22; Length 236303;  
Best Local Similarity 78.2%; Pred. No. 8.1e-68;  
Matches 419; Conservative 0; Mismatches 105; Indels 12; Gaps 3;  
QY 52 CGACTTCATGATAAACTTCAGTGGATGAGGAGCTGCTCTTAT-GATGAACAAGAAGG 110  
DB 107579 CAATCTCATGATAAACTGGAAGGATGAGGAGTTCTTCTTATGATGACCAAGAAG 107638  
QY 111 TCGTTTCTTGAATGGAATCTACTCTGTTGAAGATGCTGTGAACA-----TTGTT 161  
DB 107639 TAGTTTCTTGAATGGAATCTACTCTGTTGAAGATGCTGTGAACAATCTTACCAGAT 107698  
QY 162 GAAATGACAAAGAATTTACATAGTGTAGTATGATGAAGCAGTAGCAGGA 221  
DB 107699 AAAAGCAGCAACACGGATTTAGACTATTACATAAATTTAGTTGATAAGCAACAGCAGG 107758  
QY 222 TTCGAGAGATCGATCCCAATTTCAAAATAAGTTCTTCTGTGGTAAATGCTATCAAT 281  
DB 107759 TTTTAGAGGATTCCTCCAGTTTGAACAAGTTCTACTGTAAGTAAATGCTATCAAC 107818  
QY 282 GCGTCTGATGCTACAGAGAAATCTATCATGAAGGAGAGTCAATTCATGTTGGCAACT 341  
DB 107819 AGCATACATGCTACAAAGAAATCTTTCATGAAGGAAGAGTTAGTCAATGTGGCAACT 107878  
QY 342 TCATTTGTTGCTGATTTTAAAGAAATGTGAGGACCCCACTTCAACAACCATGACC 401  
DB 107879 TCATTTGTTGTTATTTAAGAAATTTGCCACAGCCCACTTCAACAACCATGACC 107938  
QY 402 CTGATCAGTACAGGACCATCCACATTTGAGCGGAGAACCTCCAGCAGTAAAGATTATGA 461  
DB 107939 CTTATCAGTACAGGACCATCCACATTTGAGCGGAGAACCTCCAGCAGTAAAGATTATGA 107998  
QY 462 TTCCTAAGGATCAGATGAACATTTAGCAATTTTAAAGCAATAAGTATTTTACGTAA 521  
DB 107999 TTTACTGAAGCTCAGATGATGCTGATGAGTATTTT--AGCAATAAGTATTTTAAATTA 108056  
QY 522 GATATGATGTTATTTTATGAGCAATATCTATTTATGATGCTATTTTATAGACTCCAGTA 577  
DB 108057 GGAATGTAATTTTATTTAGACATAATCTATTTGATATCTTAAATAGACTATGCA 108112

RESULT 8  
AAS98600/c

ID AAS98600 standard; DNA; 140036 BP.  
XX  
AC AAS98600;  
XX  
DT 12-MAR-2002 (first entry)  
XX  
DE Human genomic DNA for PHIP/NDRP.  
XX  
XX  
KW PHIP: Pleckstrin homology domain-interacting protein; NDRP; ds;  
KW neuronal differentiation-related protein; insulin receptor substrate;  
KW IRS; signal transducer and activator of transcription; STAT;  
KW transgenic animal; diabetes mellitus type 2; hyperglycaemia;  
KW myotonic muscular dystrophy; acanthosis; nigricans; retinopathy;  
KW nephropathy; arteriosclerosis; peripheral arterial disease; cancer;  
KW adenocarcinoma; leukaemia; breast cancer; prostate cancer; colon cancer;  
KW ovarian cancer; autoimmune disease; inflammation; immunodeficiency.  
XX  
OS Homo sapiens.  
XX  
PN WO200185785-A2.  
XX  
PN 15-NOV-2001.  
XX  
PF 10-MAY-2001; 2001WO-CA00673.  
XX  
PR 11-MAY-2000; 2000US-203561P.  
XX  
XX (ROZA/) ROZAKIS-ADCOCK M.  
PA (FARH/) FARHANG-FALLAH J.  
PA (CHEN/) CHENG A.  
PI Rozakis-Adcock M, Farhang-Fallah J, Cheng A;  
XX  
DR WPI: 2002-041586/05.  
XX  
PT Novel Pleckstrin homology domain interacting protein recruiting  
PT proteins of insulin receptor substrate family, and signal transducer  
PT and activator of transcription factors to their receptors, useful to  
PT treat diabetes  
XX  
PS Disclosure; Page 99-133; 139pp; English.  
XX  
CC The invention relates to an isolated Pleckstrin homology domain  
CC interacting protein (PHIP) that recruits proteins of the insulin  
CC receptor substrate (IRS) family, and signal transducer and activator of  
CC transcription (STAT) transcription factors, to receptors that interact  
CC with and phosphorylate the proteins and STAT transcription factors,  
CC the nucleic acid encoding PHIP (nPHIP), a nucleic acid which binds to  
CC nPHIP or regions of it, analogues, fragments or allelic variants of PHIP  
CC or nPHIP, a nucleic acid sequence having substantial sequence identity or  
CC sequence similarity with a nucleic acid sequence fully defined in  
CC human neuronal differentiation-related protein (NDRP) nucleic acid  
CC sequence or its exons as given in the specification, expression  
CC vectors and host cells expressing the nucleic acids, anti-PHP antibodies,  
CC and a transgenic animal not already expressing PHIP. The nucleic  
CC acids, proteins and antibodies are useful for diagnosis and treatment of  
CC a condition associated with an insulin receptor (e.g. diabetes mellitus  
CC type 2, hyperglycaemia, myotonic muscular dystrophy, acanthosis,  
CC nigricans, retinopathy, nephropathy, arteriosclerosis, peripheral  
CC arterial disease) or cancer (e.g. adenocarcinoma, leukaemia, breast  
CC cancer, prostate cancer, colon cancer, ovarian cancer and many others  
CC given in the specification). autoimmune disease, inflammation and  
CC immunodeficiency. The protein is also useful for discovering  
CC or testing compounds which may be either enhancers or inhibitors of PHIP  
CC function. The present sequence is genomic DNA encoding PHIP and NDRP.  
XX  
SQ Sequence 140036 BP; 48500 A; 25843 C; 22499 G; 43194 T; 0 other;  
Query Match 41.2%; Score 309.4; DB 24; Length 140036;  
Best Local Similarity 73.4%; Pred. No. 9.7e-67;  
Matches 481; Conservative 0; Mismatches 156; Indels 18; Gaps 6;  
QY 52 CGACTTCATGATAAACTTCAGTGGATGAGGAGCTGCTCTTAT-GATGAACAAGAAGG 110

Db	92199	CAATCTCATGATTAAACCTTAATAATGGGGGAGGAGTTGCTCTCTTATGATGAGCAAGAAG	92140
QY	111	TGCTTTCTTGAATGGGAATCTACTCCTGTGAAGATGCTGTGAACATTTGTTGAAATGACA	170
Db	92139	TAGTTCTTCTGAGATGGGAATTTACCTTGGTCAAGATGCTGTGAACATTTGTTGAATGAAA	92080
QY	171	AGAAAGAAATTTACAGTGTTCATAGAGTATTGTTGATGAAGCAGTACGAGGATTCGAGAGG	230
Db	92079	ACAAGGATGTGGACTATTACATGAACCTCAGTTGGTAAAGCAGTGGCTTCAATTTGACAGG	92020
QY	231	ATCGATTCCAAATTTCAAATAAAGTTCTTCTGTGGGTAAAATGCTATCAAAATGCGGTCGCA	290
Db	92019	GCTGACTCCAAATTTGAAATTAAGTTCTACTGTGGGTAAAATGCTGTCAAGATGATTTCCA	91960
QY	291	TGCTACAGAGAAATCTCATATGAAGGAAGAGT-CAAATTTGATGTGGCAAACATTCATTGTT	349
Db	91959	TGCTGCAGAGAAATCTCTTGTGAAAGGAAGAGTCCAAATGGATGCCAGCAGCTTCATTGTT	91900
QY	350	GTCTGATTTTAAAGAAATTTGCAGGACCAACCCCAAGCTTCAACAACCATGACCTGTATCAG	409
Db	91899	GTCTTATTTTAAAGAAATTTGTCACAGCTACCCCAACCTTCAGCAA-CTCCACCCCTGATCAG	91841
QY	410	TCAGGAGCCATCCACATTTGAGGCGCAGAACCTCCAGCAGTAAAAGATTTATGATCTCTAA	469
Db	91840	CCAAACAGCTGTACAGCATCAGGCAAGGCCCTCCACGACCAAAATATTACTTGTCTGA	91781
QY	470	AGGATCAGATGAACATTTAGCATTTTTTAAAGCAATAAGTATTTTACGTAAAGATTTATGTTATGC	91726
Db	91780	AGTTTCAGATAATTAGCA- - - - -TTTITTAGCAATAAGTATTTTAAAGTAAATGATATGC	91726
QY	530	TGCTATTTTITTAGGCATAATGCTATTTATGATTTTAATAGCTCCAGTATTTGTTAAACAT	589
Db	91725	CCATTTTITTAAGACATAATACTATTGCACACTTAATACACTGCAGCAATAATGTTAAACGT	91666
QY	590	AACT-TTAAATGCACGTGGAGATAAAAGTA- - - - -TTTGCCTTTTATGATATTTG	639
Db	91665	AACTCTTACATGCACITGGAAACCAAAAATTTGTTGGTTGCTTCTGTTGGAACTACTTG	91606
QY	640	CTTTATGAGTAGTGTGTAATGAAACTACATTTATCTTTGGGTACACCTGTAT	694
Db	91605	CTTAATGTCAGTGTCTGGAACCTGAGCTGGGTATCTAGGAAGTATGACTATAT	91551
RESULT 9			
ABK49586			
ID	ABK49586 standard; DNA; 143306 BP.		
XX			
AC	ABK49586;		
XX			
DT	15-JUL-2002 (first entry)		
XX			
DE	Human transporter protein gene.		
XX			
KW	Human; ds; gene; transporter; transgenic; transporter mediated disease;		
KW	drug screening; pharmacogenomic analysis; chromosome 18; SNP;		
KW	single nucleotide polymorphism.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	3000..140308	
FT		/*tag= a	
FT		/product= "Transporter"	
FT	exon	3000..3051	
FT		/*tag= b	
FT		/number= 1	
FT	intron	3052..62251	
FT		/*tag= c	
FT		/number= 1	
FT	exon	62251..62601	
FT		/*tag= d	
FT		/number= 2	

FT	variation	/*tag= ac	/standard_name= "Single-nucleotide polymorphism"
FT		replace (11470,T)	replace (11470,T)
FT		/*tag= ad	/*tag= ad
FT	variation	/standard_name= "Single-nucleotide polymorphism"	/standard_name= "Single-nucleotide polymorphism"
FT		replace (11534,C)	replace (11534,C)
FT		/*tag= ae	/*tag= ae
FT	variation	/standard_name= "Single-nucleotide polymorphism"	/standard_name= "Single-nucleotide polymorphism"
FT		replace (11540,C)	replace (11540,C)
FT		/*tag= af	/*tag= af
FT	variation	/standard_name= "Single-nucleotide polymorphism"	/standard_name= "Single-nucleotide polymorphism"
FT		replace (11557,G)	replace (11557,G)
FT		/*tag= ag	/*tag= ag
FT	variation	/standard_name= "Single-nucleotide polymorphism"	/standard_name= "Single-nucleotide polymorphism"
FT		replace (11708,C)	replace (11708,C)
FT		/*tag= ah	/*tag= ah
FT	variation	/standard_name= "Single-nucleotide polymorphism"	/standard_name= "Single-nucleotide polymorphism"
FT		replace (11953,T)	replace (11953,T)
FT		/*tag= ai	/*tag= ai
FT	variation	/standard_name= "Single-nucleotide polymorphism"	/standard_name= "Single-nucleotide polymorphism"
FT		replace (11967,C)	replace (11967,C)
FT		/*tag= aj	/*tag= aj
FT	variation	/standard_name= "Single-nucleotide polymorphism"	/standard_name= "Single-nucleotide polymorphism"
FT		replace (12172,C)	replace (12172,C)
FT		/*tag= ak	/*tag= ak
FT	variation	/standard_name= "Single-nucleotide polymorphism"	/standard_name= "Single-nucleotide polymorphism"
FT		replace (12188,A)	replace (12188,A)
FT		/*tag= al	/*tag= al
FT	variation	/standard_name= "Single-nucleotide polymorphism"	/standard_name= "Single-nucleotide polymorphism"
FT		replace (13471..13472,TTT)	replace (13471..13472,TTT)
FT		/*tag= am	/*tag= am
FT	variation	/standard_name= "Single-nucleotide polymorphism"	/standard_name= "Single-nucleotide polymorphism"
FT		replace (13781,C)	replace (13781,C)
FT		/*tag= an	/*tag= an
FT	variation	/standard_name= "Single-nucleotide polymorphism"	/standard_name= "Single-nucleotide polymorphism"
FT		replace (14031,G)	replace (14031,G)
FT		/*tag= ao	/*tag= ao
FT	variation	/standard_name= "Single-nucleotide polymorphism"	/standard_name= "Single-nucleotide polymorphism"
FT		replace (14110,C)	replace (14110,C)
FT		/*tag= ap	/*tag= ap
FT	variation	/standard_name= "Single-nucleotide polymorphism"	/standard_name= "Single-nucleotide polymorphism"
FT		replace (14248,G or T)	replace (14248,G or T)
FT		/*tag= aq	/*tag= aq
FT	variation	/standard_name= "Single-nucleotide polymorphism"	/standard_name= "Single-nucleotide polymorphism"
FT		replace (14676,C or A)	replace (14676,C or A)
FT		/*tag= ar	/*tag= ar
FT	variation	/standard_name= "Single-nucleotide polymorphism"	/standard_name= "Single-nucleotide polymorphism"
FT		replace (14981,T)	replace (14981,T)
FT		/*tag= as	/*tag= as
FT	variation	/standard_name= "Single-nucleotide polymorphism"	/standard_name= "Single-nucleotide polymorphism"
FT		replace (16365,G)	replace (16365,G)
FT		/*tag= at	/*tag= at
FT	variation	/standard_name= "Single-nucleotide polymorphism"	/standard_name= "Single-nucleotide polymorphism"
FT		replace (16827,A)	replace (16827,A)
FT		/*tag= au	/*tag= au
FT	variation	/standard_name= "Single-nucleotide polymorphism"	/standard_name= "Single-nucleotide polymorphism"
FT		replace (17599,T)	replace (17599,T)
FT		/*tag= av	/*tag= av
FT	variation	/standard_name= "Single-nucleotide polymorphism"	/standard_name= "Single-nucleotide polymorphism"
FT		replace (18093,G)	replace (18093,G)
FT		/*tag= aw	/*tag= aw
FT	variation	/standard_name= "Single-nucleotide polymorphism"	/standard_name= "Single-nucleotide polymorphism"
FT		replace (18226,C)	replace (18226,C)
FT		/*tag= ax	/*tag= ax
FT	variation	/standard_name= "Single-nucleotide polymorphism"	/standard_name= "Single-nucleotide polymorphism"
FT		replace (18467..18468,GTT or GGT)	replace (18467..18468,GTT or GGT)
FT		/*tag= ay	/*tag= ay
FT	variation	/standard_name= "Single-nucleotide polymorphism"	/standard_name= "Single-nucleotide polymorphism"
FT		replace (18511..18512,TAC or TTC)	replace (18511..18512,TAC or TTC)
FT		/*tag= az	/*tag= az
FT	variation	/standard_name= "Single-nucleotide polymorphism"	/standard_name= "Single-nucleotide polymorphism"
FT		replace (19107,A)	replace (19107,A)
FT		/*tag= ba	/*tag= ba

FT	variation	/standard_name= "Single-nucleotide polymorphism" replace (19223,G) /*tag= bb	40.7%; Score 305.8; DB 24; Length 143306; Best Local Similarity 74.0%; Pred. No. 7.7e-66; Matches 485; Conservative 0; Mismatches 152; Indels 18; Gaps
FT	variation	/standard_name= "Single-nucleotide polymorphism" replace (20102,C) /*tag= bc	
FT	variation	/standard_name= "Single-nucleotide polymorphism" replace (21946,T) /*tag= bd	
FT	variation	/standard_name= "Single-nucleotide polymorphism" replace (22240,C) /*tag= be	
FT	variation	/standard_name= "Single-nucleotide polymorphism" replace (22679,G) /*tag= bf	
FT	variation	/standard_name= "Single-nucleotide polymorphism" replace (23178,C or G) /*tag= bg	
FT	variation	/standard_name= "Single-nucleotide polymorphism" replace (23370,G) /*tag= bh	
FT	variation	/standard_name= "Single-nucleotide polymorphism" replace (23583,G) /*tag= bi	
FT	variation	/standard_name= "Single-nucleotide polymorphism" replace (24159,A) /*tag= bj	
QY			Query Match 40.7%; Score 305.8; DB 24; Length 143306; Best Local Similarity 74.0%; Pred. No. 7.7e-66; Matches 485; Conservative 0; Mismatches 152; Indels 18; Gaps
Db		47 ATGGTCGACATTCATGATAAACTTCAGTGGATGAGGAGTGCCTCTTAT -GATGAACAAA 105 	
Db		88548 ATGCTGCATCTCATGATAAACTTGAATCAATTAGGAGCTCTTCTATGATGAACAAA 88607 	
QY		106 GAAGGTGGTTCTTGAATGAATCTACT - - - CCTGGTGAAGATGCTGTGAACATTGTT 161 	
Db		88608 GAAGTGATTCTTGAGATGGAATCTGCTCGCTGGTAAAGATGCTGTGAACATTGTT 88667 	
QY		162 GAAATGACAAAGAAGATTTACAGTGTTCATAGAGTTAGTTGATGAAGCAGTAGCAGGA 221 	
Db		88668 GAAATGACAAAGAAGATTCAGATATTTATCACTTACTAGTGTAAAGCAGTGGCAGGG 88727 	
QY		222 TTGAGAGGATCGATTCCAAATTCAAAATAAGTTCTTCTGGTAAAAATGCTATCAAT 281 	
Db		88728 TTTGAGAAGATTGATTCCAATTTGAAAA -AAGTTCTACTGTAGTAAATGCTATCTAAC 88786 	
QY		282 GCGTCGCGATCTACAGAGAAATCTATCATGAAGAGAGATCAATTGATGTGGCAACT 341 	
Db		88787 AGCATTTGCATGTTACAGAGAA - - - ATTGTGAAGAGAGAGATCAATGGATGTGCCAACT 88842 	
QY		342 TCATTTGTTGCTGTTATTTAAGAAATTTGTCAGGACCAACCCCTTCAACAACCATGACC 401 	
Db		88843 TCCTTTGTTGCTTATTTTCAGAAATTTGTCAGCACCACCCCGCCCTTCAACAGTCAACGCC 88902 	
QY		402 CTGATCAGTCAGGAGCCATCCCATTTGAGGGGAGAACCTCCAGCAGTAAAAAGATTATGA 461 	
Db		88903 CTGATCAGTCAGTAGTCATCAACATCAAGAGAGACCCTCCGCGAGAGCATTATGA 88962 	
QY		462 TTCTCTAAGGATCAGATGAACATTAGCATTTTTTTAAGCAATAAAGTATTTTTACGTAA 521 	
Db		88963 CTTGTGAAACCTCAGATTATTTAGAAATTTTGGTAAATAAGTATTTTCAAGTTAAAG 89022 	
QY		522 GATATGTATGTTATTTTTTAGGCATATGCTATTATGCAATTTAATAGACTCCAGTATATT 581 	
Db		89023 AGTATATAT -ACATTTTTTAGACCTAATGCTATTGTACACTCAATAGACTGTAGTAGT 89081 	
QY		582 GTAAACATAAC -TTTAAATGCATGGGAGATAAAAGTATTTCCTCTTTTATCATATTTGC 640 	
Db		89082 GTAAACATAACITTTATATGCATCGGGAACAAA - - - - -AAGAAATGTGTGACTCAT 89135 	
QY		641 TTTATGTCAGTAGTCGTGAATGGAACTACATTAATCTCTGGGTACACCTGTATA 695 	

Query Match	40.7%;	Score 305.8;	DB 24;	Length 143306;
Best Local Similarity	74.0%;	Pred. No. 7.7e-66;		
Matches 485;	Conservative 0;	Mismatches 152;	Indels 18;	Gaps
47	ATGGTCGACTTCATGATAAACTTCAGTGGATGAGAGCTGCCTCTTAT-GATGAACAAA	105		
88548	ATGCTGCATCTCATGATAAACTTGAATCAATTAGAGAGCTTCTTTATGGATGAACAAA	88607		
106	GAAGTGTGTTCTTCGAAATGGAATCTACT- - -CCTGGTGAAGATGCTGTGACATTTGTT	161		
88608	GAAGTGAATTTCTGAGATGGAATCTGCTCCGCTCGTGAAGATGCTGTGAACATTTGT	88657		
162	GAATGACAAGAAGAATTTACAGTGTTCACATAGAGTTAGTTGATGAAGCAGTAGCAGGA	221		
88668	GAATGACAACAAGAATTCAGAAATTTATATCAACTTAGTTGATTAAGCAGTAGTGCACGG	88727		
222	TCGAGAGATGCGATTCCAAATTTCAAAATAAGTTCTTCTGTGGTAAATGCTATCAAAAT	281		
88728	TTTGAGAAGATTGATTCCAATTTGAAAA- AAGTTCCTACTGTAGGTAAATGCTATCTAAC	88786		
282	GCCTGCGCATGCTACAGAGAAATCTATCATCAAGGAAGAGTCAATTGATGTGGCAAACT	341		
88787	AGCATTGCAATGTTACAGAGAA- - -ATTGTGAAAGGAAGAGTCAATGGATGTGGCCAACT	88842		
342	TCATTGTTGTCGTATTTTAAAGAAATGTGCAGGACACCCCAACCTTCACAACCATGACC	401		
88843	TCTTGTTGTCCTATTTTCAAGAAATGTGCAGCCACCCCAACCTTCACAAGTCAACGCC	88902		
402	CTGATCAGTCAGGAGCCATCCACATTGAGGCGGAGAACTCCAGCAGTAAAAAGATTATGA	461		
88903	CTGATCAGTCAGTAGTCATCAACATCAAGGAAGACCCCTCGCCAGCGAAGCAATTATGA	88962		
462	TTCCTAAGAGATCAGATGAACATTAGCATTTTTTTTAAAGCAATAAAGATTTTTTACGTAA	521		
88963	CTTGCGAAACTCAGATTTATTAGAAATTTTGGTAAATAAGATTTTTCAGGTTAAGG	89022		
522	GATATGTATGTTATTTTTTAGGCATAATGCTATTATGCAATTTAATAGACTCCAGTATATT	581		
89023	AGTATATAT-ACATTTTTTAGACCTAATGCTATTGTACACTCAATAGACTGTAGTAGAT	89081		
582	GTAACACATAAC-TTTAAATGCACCTGGGAGATAAAAGTATTTGCTCTTTTATGATATTGC	640		
89082	GTAACACATACTTTTATATGCACCTGGGAACAAA- - - - -AGAAAAATGTGTGACATCAT	89135		
641	TTTATTGCGTAGTCTGTAAATGGAACATACATTATCTCTGGGTACACCTGTGATA	695		

Db	89136	TTTTATTGTGGTGCTCAAGTCTAACTCAACAATGCTCTGACGTATGCCTGTATA	89190
RESULT 10			
ABAL8218			
ID	ABAL8218	standard; DNA; 5022 BP.	
XX	XX		
AC	ABAL8218;		
XX	XX		
DT	23-JAN-2002	(first entry)	
XX	XX		
DE	Human nervous system related polynucleotide SEQ ID NO 10549.		
XX	XX		
KW	Human; neurotropic; neuroprotective; cytostatic; dermatological; virucide;		
KW	immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;		
KW	antiparkinsonian; antiskloking; antianemic; antithratic; cancer;		
KW	antirhematic; hepatotropic; cerebroprotective; antiinflammatory;		
KW	antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;		
KW	antiparasitic; cardiant; immune disorder; cardiovascular disorder;		
KW	neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.		
OS	Homo sapiens.		
XX	XX		
PN	WO200159063-A2.		
XX	XX		
PD	16-AUG-2001.		
XX	XX		
PF	17-JAN-2001; 2001WO-US01334.		
XX	XX		
XX	31-JAN-2000; 2000US-0179065.		
PR	04-FEB-2000; 2000US-0180628.		
PR	24-FEB-2000; 2000US-0184564.		
PR	02-MAR-2000; 2000US-0186350.		
PR	16-MAR-2000; 2000US-0189874.		
PR	17-MAR-2000; 2000US-0190076.		
PR	18-APR-2000; 2000US-0198123.		
PR	19-MAY-2000; 2000US-0205515.		
PR	07-JUN-2000; 2000US-0209467.		
PR	28-JUN-2000; 2000US-0214886.		
PR	30-JUN-2000; 2000US-0215135.		
PR	07-JUL-2000; 2000US-0215647.		
PR	07-JUL-2000; 2000US-0216680.		
PR	11-JUL-2000; 2000US-0217487.		
PR	11-JUL-2000; 2000US-0217496.		
PR	14-JUL-2000; 2000US-0218290.		
PR	26-JUL-2000; 2000US-0220963.		
PR	26-JUL-2000; 2000US-0220964.		
PR	14-AUG-2000; 2000US-0224518.		
PR	14-AUG-2000; 2000US-0224519.		
PR	14-AUG-2000; 2000US-0225213.		
PR	14-AUG-2000; 2000US-0225214.		
PR	14-AUG-2000; 2000US-0225266.		
PR	14-AUG-2000; 2000US-0225267.		
PR	14-AUG-2000; 2000US-0225268.		
PR	14-AUG-2000; 2000US-0225270.		
PR	14-AUG-2000; 2000US-0225447.		
PR	14-AUG-2000; 2000US-0225757.		
PR	14-AUG-2000; 2000US-0225758.		
PR	14-AUG-2000; 2000US-0225759.		
PR	18-AUG-2000; 2000US-0226279.		
PR	22-AUG-2000; 2000US-0226681.		
PR	22-AUG-2000; 2000US-0226868.		
PR	22-AUG-2000; 2000US-0227182.		
PR	23-AUG-2000; 2000US-0227009.		
PR	30-AUG-2000; 2000US-0228924.		
PR	01-SEP-2000; 2000US-0229287.		
PR	01-SEP-2000; 2000US-0229343.		
PR	01-SEP-2000; 2000US-0229344.		
PR	01-SEP-2000; 2000US-0229345.		
PR	05-SEP-2000; 2000US-0229509.		
PR	05-SEP-2000; 2000US-0229513.		
PR	06-SEP-2000; 2000US-0230437.		
PR	06-SEP-2000; 2000US-0230438.		
PR	08-SEP-2000; 2000US-0231242.		
PR	08-SEP-2000; 2000US-0231243.		
PR	08-SEP-2000; 2000US-0231244.		
PR	08-SEP-2000; 2000US-0231413.		
PR	08-SEP-2000; 2000US-0231414.		
PR	08-SEP-2000; 2000US-0232080.		
PR	08-SEP-2000; 2000US-0232081.		
PR	12-SEP-2000; 2000US-0231968.		
PR	14-SEP-2000; 2000US-0232397.		
PR	14-SEP-2000; 2000US-0232398.		
PR	14-SEP-2000; 2000US-0232399.		
PR	14-SEP-2000; 2000US-0232400.		
PR	14-SEP-2000; 2000US-0232401.		
PR	14-SEP-2000; 2000US-0233063.		
PR	14-SEP-2000; 2000US-0233064.		
PR	21-SEP-2000; 2000US-0233065.		
PR	21-SEP-2000; 2000US-0234223.		
PR	21-SEP-2000; 2000US-0234274.		
PR	25-SEP-2000; 2000US-0234937.		
PR	25-SEP-2000; 2000US-0234938.		
PR	26-SEP-2000; 2000US-0235484.		
PR	27-SEP-2000; 2000US-0235834.		
PR	27-SEP-2000; 2000US-0235836.		
PR	29-SEP-2000; 2000US-023		



PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 01-DEC-2000; 2000US-0251160.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 06-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 DR WPI; 2001-541565/60.  
 XX  
 XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating nervous system  
 PT cancers and metastases -  
 XX  
 PS Disclosure; SEQ ID NO 10549; 1701pp + Sequence Listing; English.  
 XX  
 CC The invention relates to novel genes (ABA11004-ABA21534) and proteins  
 CC (ABBI4678-ABBI8001) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;  
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
 CC and parasitic infections.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).  
 XX  
 SQ Sequence 6022 BP; 1762 A; 1163 C; 1219 G; 1878 T; 0 other;

Query Match 40.2%; Score 302.2; DB 22; Length 6022;  
 Best Local Similarity 75.2%; Pred. No. 2.5e-65;  
 Matches 482; Conservative 0; Mismatches 148; Indels 11; Gaps 8;

QY 60 TGATAAACTTCAGTGGATGAGGAGGCTCTTATG-ATGAACAAAGAGGTGTTCT 118  
 DB 2031 TCATAAACTTGAATGATGAGGAGGCTCTTATGATGAGCAAGAGGAGGTTCT 2090  
 QY 119 TGAATGGATCTACTCTGCTGAGAGTCTGCTGAACATTTGTTAAATGACAAAGAA 178  
 DB 2091 TGAGTAGAATCTAGTCTGCTGAGAGTCTGCTGAGCAATTTTAACTGGCAACGAGGA 2150  
 QY 179 TTACAGTGTACATAGATTAGTTGATGAGCAGTAGCAGAGTTCGAGAGATCGATTC 238  
 DB 2151 TTAGATATGACTTAACTTAGTTGAAAGAGGAGCAGC-GGGTCTGAGAGGATGACTA 2209  
 QY 239 CAATTTCAAAATAGTCTTCTGGTGAATAATGCTCAATGCGTGGATGCTACAG 298  
 DB 2210 CAGTTTGAAGAAGTCTTACTGTGGATAAATACTATCACACACACTGCAITGCTACAG 2269

QY 299 AGAATCTATCATGAAAGGAGAGTCATTTGATGTGGCAAACTTCATTTGTTGCTATTT 358  
 DB 2270 AGAAGTTTTGGTGAARAGAGAGTCAGTTGATGAGGAGAGCTTTACTGTTGCTTATTT 2329  
 QY 359 TAAGAATTTCTCAGGACCAACCCCAACCTTCAACAAACCATGACCTGATCAGTCAGAGCC 418  
 DB 2330 TATGCAATTG-CACAGCCACCTCATTTAGCAACCTCCATCTCTGATCAGTCAGCAGCC 2388  
 QY 419 ATC-CACATTGAGGAGAGACCTCCAGCAGTAAAGATTTATGATCTCTAAAGGATCAG 477  
 DB 2389 ATCAGCATGGAGGCAAGACCTCCATCAGCAAAAGATTAAGACTCTCTGAAGGCTCAG 2448  
 QY 478 ATGAACATTAGCATTTTTTAAGCAATAAAGTATTTTACGTAAGATATGATGTTATT- 536  
 DB 2449 ATGATCATTAACA-TTTTGTAGCAATAAAGTATTTTAAATTAAGTGTAGTACATTGTTA 2507  
 QY 537 TTTTAGGCATAATGCTATTATGCAATTTAATAGACTCCAGTATATTGTAACAT---AACT 593  
 DB 2508 TTTTACACATAATGCTAT--TGCACCTAATAGACTACATGTAATGTAATATATAACTT 2565  
 QY 594 TTAATGCACTGGGAGATAAAAGTATTTGCTCTTTTATGATATTTGCTTTTATGCGAGTAG 653  
 DB 2566 TTATATGCCACCGAGAAACCAAAATTTGTGACTCAGTATTTTATGTTGGTAG 2625  
 QY 654 TCTGTAATGGAACACTACATTATCTCTTGGGTACACCTGTAT 694  
 DB 2626 TCTGGAACCAACCAACCAATATCTCCAGGTAGGCTGTAT 2666

RESULT 11  
 ABL65786/c  
 ID ABL65786 standard; DNA; 460 BP.  
 XX  
 AC ABL65786;  
 XX  
 DT 15-MAY-2002 (first entry)  
 XX  
 DE Lung cancer related gene sequence SEQ ID NO:4123.  
 XX  
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
 KW cytosatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
 KW gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200194629-A2.  
 XX  
 PD 13-DEC-2001.  
 XX  
 PF 30-MAY-2001; 2001WO-US10838.  
 XX  
 PR 05-JUN-2000; 2000US-209473P.  
 PR 05-JUN-2000; 2000US-209531P.  
 PR 18-SEP-2000; 2000US-233133P.  
 PR 18-SEP-2000; 2000US-233617P.  
 PR 20-SEP-2000; 2000US-234009P.  
 PR 20-SEP-2000; 2000US-234034P.  
 PR 20-SEP-2000; 2000US-234052P.  
 PR 22-SEP-2000; 2000US-234509P.  
 PR 22-SEP-2000; 2000US-234567P.  
 PR 22-SEP-2000; 2000US-234923P.  
 PR 25-SEP-2000; 2000US-234924P.  
 PR 25-SEP-2000; 2000US-235077P.  
 PR 25-SEP-2000; 2000US-235082P.  
 PR 25-SEP-2000; 2000US-235134P.  
 PR 25-SEP-2000; 2000US-235280P.  
 PR 26-SEP-2000; 2000US-235637P.  
 PR 26-SEP-2000; 2000US-235638P.  
 PR 27-SEP-2000; 2000US-235711P.  
 PR 27-SEP-2000; 2000US-235720P.  
 PR 27-SEP-2000; 2000US-235840P.

27-SEP-2000; 2000US-235863P.  
28-SEP-2000; 2000US-236028P.  
28-SEP-2000; 2000US-236032P.  
28-SEP-2000; 2000US-236033P.  
28-SEP-2000; 2000US-236034P.  
28-SEP-2000; 2000US-236109P.  
28-SEP-2000; 2000US-236111P.  
28-SEP-2000; 2000US-236842P.  
29-SEP-2000; 2000US-236891P.  
29-SEP-2000; 2000US-237172P.  
02-OCT-2000; 2000US-237173P.  
02-OCT-2000; 2000US-237278P.  
02-OCT-2000; 2000US-237294P.  
02-OCT-2000; 2000US-237295P.  
02-OCT-2000; 2000US-237316P.  
03-OCT-2000; 2000US-237425P.  
03-OCT-2000; 2000US-237598P.  
03-OCT-2000; 2000US-237604P.  
03-OCT-2000; 2000US-237606P.  
03-OCT-2000; 2000US-237608P.  
01-NOV-2000; 2000US-244867P.  
01-NOV-2000; 2000US-245084P.  
XX  
XX  
(AVAL-) AVALON PHARM.  
PA  
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
PI Soppet DR, Weaver Z;  
XX  
XX  
WPI; 2002-188264/24.  
XX  
XX  
Screening for anti-neoplastic agent involves exposing cells to a  
PT chemical agent to be tested for anti-neoplastic activity, and  
PT determining a change in expression of a gene of a signature gene set -  
XX  
XX  
Claim 1; SEQ ID 4123; 44pp; English.  
XX  
XX  
The present invention describes a method (M1) for screening for an  
CC anti-neoplastic agent. The method involves exposing cells to a chemical  
CC agent to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 847 sequences (given in ABU61664  
CC to ABL70110), or is at least 9% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening  
CC an anti-neoplastic agent, and can be used for producing a product which  
CC is the data collected with respect to the anti-neoplastic agent as a  
CC result of M1, and the data is sufficient to convey the chemical  
CC structure and/or properties of the agent. M1 can be used in the  
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
CC carcinoma, papillary carcinoma and Wilms' tumour.  
XX  
SQ Sequence 460 BP; 114 A; 92 C; 94 G; 160 T; 0 other;

Query Match 40.1%; Score 301; DB 24; Length 460;  
Best Local Similarity 81.9%; Pred. No. 2.4e-65;  
Matches 371; Conservative 0; Mismatches 80; Indels 2; Gaps 2;  
QY 65 AAACCTCAGTGAGGAGCGCTCTTAT-GATCAACAAGAGAGGCTTCTTCAAA 123  
DB 460 AAACCTTACAGATGAGAGTGTCTTATGATCGCAAGAAAGTGTCTTCTGAGA 401  
QY 124 TGGAACTCTACTCTCTGGTGAAGATGCTGTGAACATTTGTAATGACAAGAAAGATTTAC 183  
DB 400 TGGAACTCTACACCTTGGTGAAGATGCTGTGAACATTTGTAATGACAAGAAAGATTTAG 341  
QY 184 AGTGTATACATAGATTAGTTGATGAAGCAGTACAGGATTCGAGAGATCGATTCCAATT 243  
DB 340 AATATTACATAAATCTAGTTGATAAAGCAGCAGGAGGTTTGAGAGGATTCATCCAATT 281  
QY 244 TCAAAATAAGTCTCTCTGGTGAAGTCTATCAATGCGCTCGCATGCTACAGAGAA 303

Db 280 TTGAAGAAGCTCTACTGTGGCAAAATGCTATCAAAACAGCATCGCATGCTACAGAGAA 221  
QY 304 TCTATCATGAAGGAAGAGTCAATTGATGTGCGCAACTTCATTGTTGTCGTATTTAAGA 363  
Db 220 TCTTTCATGAAGGAAGAGTCAACTCATCGGAAAGCTTCACCGATGTCTTATTTAGGA 161  
QY 364 AATTGTCTAGGACACCCCAACCTTCAACAACCATGATGATCAGTCAGGAGCCATCCA 423  
Db 160 AATTGCCACAGCCACCCCAACCTTCAGGAGCCACCACTGACCATGACCAACCATCAA 101  
QY 424 CATTGAGCGAGAACCTCCAGCAGTAAAGAGATTGATTCTCTAAAGGATCAGATGAAC 483  
Db 100 CCTCGAGCAAGACCTCCACAGCAAAAGAGTACGACTCACCAGGAGCTCAGATTGAT 41  
QY 484 ATTAGCATTTTTTAAAGCAATAAAGTATTTTA 516  
Db 40 AATCAGCATTTTTT-AGCAATAAAGTATTTTA 9  
RESULT 12  
AAL05284/c  
ID AAL05284 standard; DNA; 4503 BP.  
XX  
AC AAL05284;  
XX  
AC AAL05284;  
DT 21-NOV-2001 (first entry)  
XX  
XX Human reproductive system related antigen DNA SEQ ID NO: 7972.  
DE Human reproductive system related antigen; reproductive system disorder;  
KW Human; reproductive system related antigen; cancer; gene therapy; ds.  
KW  
XX  
OS Homo sapiens.  
XX  
PN WO200155320-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01339.  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180828.  
PR 24-FEB-2000; 2000US-0184564.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226581.



Db	524	TGCTACAGGGAAAACTTTCTGTGAAAGGAAGAACTCAACTGATGTGGCAAAATTTCATTGCTCT	465
Qy	351	TCGTATTTTAAAGAAATTTGTCAGGACCACCCCAACCTTTCAACAACCATGACCCGTACAGT	410
Db	464	TCATTATTTTAAAGAAATTTGCCACATGCAATCCAACTTTCAGTATACCCAGCCT-ATCAGT	406
Qy	411	CAGGAGGCATCCACATTGAGGCAGAGAACCCTCCAGCAGTAAAAAGATTTATGATTCCTCAA	470
Db	405	CAGCAGGCATCAACATTGAGGCAGACCCTCCACTGGCAAAAAGATTTACGACCTGCTGAA	346
Qy	471	GGATCAGATGAACATTTAGCATTTTTTTTAAAGCAATAAAGTATTTTTAGTAAAGATATGAT	530
Db	345	GGTCGATGATGCTTAACATTTTTTAGCAGTAAAGCGTTTTCTAAATTAAGGTATGCTC	286
Qy	531	GTT-----ATTMTTAGGCATAAGTCTATTATGCATTTAATAGACTCCACGATATATTG	583
Db	285	CTTTAAAAAAGAAAAAGAAATAAGTCTTTTCACACTTACTAGACTACAGTATAGCGT	226
Qy	584	AAACATAACTT-TAAATGCACCTGGGAGATAAAGTAT-----TTGCTCTTTTATGA	633
Db	225	AAACATAACTTCCATATGCGGTGGGAAACCAAAACAATTTGTGTGACITTCGTTATTGCAG	166
Qy	634	TATTTGCTTTTATGCAGTAGTCTGTAATGGAACACTACATATCTCTTTGGGTACACCTGTA	693
Db	165	TATTTGCTTTTGTCTGTGCTGAATCCAAATTTGCAACATCTCTAAAGATATGCCTGTA	106
Qy	694	TACAGAAAGAAATTA	709
Db	105	ATCAATATGGTAATA	90

RESULT 13

AAF21105  
ID AAF21105 standard; DNA; 143068 BP.

XX

AC AAF21105;

XX  
DT 14-MAR-2001 (first entry)

DE Human low adenosine antisense oligonucleotide related sequence #2672.

Low adenosine antisense oligonucleotide; phosphorothioate; allergy; human; airway disorder; bronchoconstriction; lung inflammation; surfactant depletion; respiratory bronchodilator; antiinflammatory; immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic; respiratory obstruction; pulmonary obstruction; impeded respiration; surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS; respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis; pulmonary hypertension; emphysema; pulmonary transplantation rejection; chronic obstructive pulmonary disease; pulmonary infection; bronchitis; cancer; ss.

Homo sapiens.

PN WO200062736-A2.

PD 26-OCT-2000.

24-MAR-2000; 2000WO-US08020.

PR 06-APR-1999; 99US-0127958.

PA (UYEC-) UNIV EAST CAROLINA.

PA (NYCE/) NYCE J W.

XX

PI Nýce JW;  
yy

XX  
DR  
WPI: 2000-679539/66.

Low adenosine (A) content antisense oligonucleotides which do not trigger adenosine receptors during metabolism, useful e.g. for treating cancers and respiratory obstructions -

[illegible]

## Query Match

Query Match 39.2%; Score 294.6; DB 21; Length 143068;  
Best Local Similarity 73.6%; Pred. No. 4.7e-63;

Matches 475; Conservative

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

QY	57	TCATGATAAAACTTCAGTGGATGAGAGCTGCCCTCTTA-TGATGAACAAAAGAAAGGTGGTT	115
Db	53544	TCATGATCAAACTTGACACAGATGAAGAGTTCGTTCTTACATATGACGAGTGAAGTGGTC	53603
QY	116	TCTTGAATGGAATCTACTCTCGTGAAGATGCTGTGAACATTTGTTGAAATGACAAGAAA	175
Db	53604	TCTTGAGATGGAATCTCCTCTCGTGAAGATGCTGTGAACACGGTTAAAAATGACAAACAAT	53663
QY	176	GAATTTACAGTGTACATAGATTAAGTTGATGAGACGAGTACGAGGATTCGAGAGGATCGA	235
Db	53664	CGATTTAGAAATTTACATAAAATTTAGTTAATAAAGCAGTGGCAGGGTTTGAGAGGATTGA	53723
QY	236	TTCCAAATTCAAAATAAGTTCCTCTGTGGGTAAAAATGCTATCAAAATGGCGTCGCATGGTA	295
Db	53724	CTCCAAATTTGAAAGAA-----GTGGGTAAATATGCTATCAAAATGCATCAATGTTGA	53775

[illegible]

QY	471	GGATCAGATGAACATTAGCATTTTTTAAAGCAATAAAGTATTTTTACGTAAAGATATGTAT	530
Db	53956	GGCTCAGATGATGGTTAGCATTTTT--AGCAATACAAATTTTTTAATTAAGGTATGCAC	54012
QY	531	GTT-ATTTTTAGCATAAATGCTATATATGCATTTAAATAGACTCCAGTATATTTGAAACAT	589
Db	54013	ATTGGTTTTCTGACATAAATACATATGCATCTTAAATAGACTACAGTATAGGATAAACAAC	54072

QY 590 AAC-TTTAAATGACTGGGAGATAAAAGTATTTGCTCTTTTATGATATTTGCTTTATGTC 648  
DB 54073 AAC-TTTTATATGCTACTGGCAACCAAAAGGTT--ATTTTGAGATATTTGCTTTACTGT 54130  
QY 649 AGTACTCTGTAATGGAACACTACATTTCTCTTGGGTACACTGTA 693  
DB 54131 GGTGCTCTGAAGCTGAACACTCAACAATCTCACCRAAGGTGTGCTGAA 54175

RESULT 14  
AAF21272  
ID AAF21272 standard; DNA; 143068 BP.  
XX AC AAF21272;  
XX AC AAF21272;  
DT 14-MAR-2001 (first entry)  
XX

DE Human low adenosine antisense oligonucleotide related sequence #2839.  
XX  
KW Low adenosine antisense oligonucleotide: phosphorothioate; allergy;  
KW human; airway disorder; bronchoconstriction; lung inflammation;  
KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;  
KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;  
KW respiratory obstruction; pulmonary obstruction; impeded respiration;  
KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;  
KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;  
KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;  
KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;  
KW cancer; ss.

OS Homo sapiens.  
XX  
XX WO200062736-A2.  
XX  
XX 26-OCT-2000.  
XX  
XX 24-MAR-2000; 2000WO-US08020.  
XX  
XX 06-APR-1999; 99US-0127958.  
XX  
XX (UYEC-) UNIV EAST CAROLINA.  
XX (NYCE/) NYCE J W.  
XX  
XX Nyce JW;  
XX  
XX WPI; 2000-679539/66.  
XX

DR Low adenosine (A) content antisense oligonucleotides which do not  
PT trigger adenosine receptors during metabolism, useful e.g. for treating  
PT cancers and respiratory obstructions -  
XX  
XX Disclosure; Page 1186-1219; 1592pp; English.  
XX

XX The present invention describes low adenosine (A) content antisense  
CC oligonucleotides and compositions (I) comprising them. In the antisense  
CC oligonucleotides the A is replaced by a 'Universal' or alternative base.  
CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,  
CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.  
CC The antisense oligonucleotides and (I) can be used to down-regulate the  
CC expression and/or activity of target polypeptides associated with  
CC lung/respiratory disorders and malignancies, such as stimulating and  
CC activating peptide factors and transmitters, transcription factors,  
CC immunoglobulins and antibodies, antibody receptors, cytokines and  
CC chemokines, endogenously produced specific and non-specific enzymes,  
CC binding proteins, adhesion molecules and their receptors, cytokine and  
CC chemokine receptors, adenosine receptors, bradykinin receptors, central  
CC nervous system (CNS) and peripheral nervous and non-nervous system  
CC receptors, CNS and peripheral nervous and non-nervous system peptide  
CC transmitters, defensins, growth factors, vasoactive peptides and  
CC receptors, binding proteins and malignancy associated proteins. The  
CC antisense oligonucleotides may be used in this way to treat disorders  
CC including respiratory obstruction (especially pulmonary obstruction

CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)  
CC and/or surfactant hypoproduction which are associated with a disease or  
CC condition selected from pulmonary vasoconstriction, inflammation,  
CC allergies, asthma, impeded respiration, respiratory distress syndrome  
CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary  
CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),  
CC pulmonary transplantation rejection, pulmonary infections, bronchitis,  
CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide  
CC fragments and antisense oligonucleotides used in the exemplification of  
CC the present invention.  
XX

SQ Sequence 143068 BP; 41194 A; 30122 C; 32403 G; 39349 T; 0 other;  
Query Match 39.2%; Score 294.6; DB 21; Length 143068;  
Best Local Similarity 73.6%; Pred. No. 4.7e-63;  
Matches 475; Conservative 0; Mismatches 149; Indels 21; Gaps 7;

QY 57 TCATGATAAACTTCAGTGATGAGGAGCTGCCTCTTA-TGATGAACAAAGAGGTGTT 115  
DB 53544 TCATGATCAAACTTGACAGATGAAGAGTGTCTTTCATATGAGCTGAAGTGTTC 53603  
QY 116 TCTTGAATGGAATCTACTCTCTGTGAAGATGCTGTGAACATTTGAAATGACAAGAA 175  
DB 53604 TCTTGAGATGGAATCTCTCTGTGAAGATGCTGTGAACAGCGTTAAATGACAACAAT 53663  
QY 176 GAATTTACAGTGTACATGATGATGATGAAGCAGTACGAGGATTCGAGAGGATCA 235  
DB 53664 CGATTTAGAAATATTACATAAAATTTAGTTAATAAGCAGTGCAGGGTTGAGAGGATGA 53723  
QY 236 TTCCAAATTTCAAAATAAGTCTCTGTGGGTAAATGCTATCAATGCGGTGCGATGTA 295  
DB 53724 TTCCAAATTTGAAAGAA-----GTGGGTAAATGCTATCAATGATCATCATCAT 53775  
QY 296 CAGAGAAATCTATCATGAAGGAGAGTCAATGATGTGCAAACTTCATTTGCTGCTA 355  
DB 53776 TGAAGAAATCTTTTGAAGGGAGAGTGCACCAAGGTGGCAATTTGATTTGCTATCTTA 53835  
QY 356 TTTTAAGAAATTTGTCAGGACCCCAACCTTCAACCAACCATGACCTGATCAGTCAGGA 415  
DB 53836 TTTTAAGAAATTTGCCACAGCCACCCAGCTTTAGCAACCAACCATGATCAGTAAGA 53895  
QY 416 GCCATCCACATTTGAGCGGAGACCTC-----CAGCAGTAAAAAGATTATGTTCTTAA 470  
DB 53896 GCCATCAACATCAAAACAAGACGCGCATCTCTTCAGCAAAAACACATGACTTGTCTGA 53955  
QY 471 GGATCAGATGAACATAGCATTTTTTAAAGCAATAAGTATTTTACGTAGATATGAT 530  
DB 53956 GGCTCAGATGATGGTAGCATTTTT---AGCAATACAAATTTTTTAATTAAGTATGCAC 54012  
QY 531 GTT-ATTTTATAGGCAATAATGCTATTATGCAATTTAATAGACTCCAGTATATTGTAACAT 589  
DB 54013 ATTGTTTTCTGCACATAATATGCTATTTGCTATTTAATAGACTACAGTATAGGATAAACAC 54072  
QY 590 AAC-TTTAAATGCACTGGGAGATAAAAGTATTTGCTCTTTTATGATATTTGCTTTATGTC 648  
DB 54073 AAC-TTTTATATGCTACTGGGAAACCAAAAGGTT--ATTTTGAGATATTTGCTTTACTGT 54130  
QY 649 AGTACTCTGTAATGGAACACTACATTTCTCTTGGGTACACTGTA 693  
DB 54131 GGTGCTCTGAAGCTGAACACTCAACAATCTCACCRAAGGTGTGCTGAA 54175

RESULT 15  
AAA34983  
ID AAA34983 standard; DNA; 143068 BP.  
XX AC AAA34983;  
XX AC AAA34983;  
DT 28-JUL-2000 (first entry)  
XX  
DE Human adenosine receptor related polynucleotide SEQ ID NO:2672.  
XX  
XX Human; adenosine receptor; low adenosine antisense oligonucleotide;



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 15:35:54 ; Search time 836.08 Seconds  
(without alignments)  
14547.424 Million cell updates/sec

Title: US-09-914-152-3\_COPY\_2000\_2750

Perfect score: 751

Sequence: 1 cctgcttggtggtcccaag.....tgatggagctggaaagtcc 751

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- EST:\*
- 1: em\_estba:\*
  - 2: em\_esthum:\*
  - 3: em\_estin:\*
  - 4: em\_estm:\*
  - 5: em\_estov:\*
  - 6: em\_estpl:\*
  - 7: em\_estro:\*
  - 8: em\_hic:\*
  - 9: gb\_est1:\*
  - 10: gb\_est2:\*
  - 11: gb\_hic:\*
  - 12: gb\_est3:\*
  - 13: gb\_est4:\*
  - 14: gb\_est5:\*
  - 15: em\_estfun:\*
  - 16: em\_estom:\*
  - 17: gb\_gss:\*
  - 18: em\_gss\_hum:\*
  - 19: em\_gss\_inv:\*
  - 20: em\_gss\_pln:\*
  - 21: em\_gss\_vit:\*
  - 22: em\_gss\_fun:\*
  - 23: em\_gss\_mam:\*
  - 24: em\_gss\_mus:\*
  - 25: em\_gss\_other:\*
  - 26: em\_gss\_pro:\*
  - 27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	309.2	41.2	515	14	BM691815
C 2	309.2	41.2	558	13	BM668477
C 3	309	41.1	465	14	N24942
4	303.6	40.4	743	17	A2519098
C 5	301.2	40.1	318	10	AW859148
C 6	301	40.1	460	9	AA608733

7	297	39.5	857	17	A0750429
8	294.2	39.2	673	17	A0312301
C 9	290.6	38.7	914	9	AL521707
10	290.4	38.7	480	14	BM703722
C 11	289.8	38.6	512	9	AL189864
C 12	288.4	38.4	542	17	A0237336
C 13	285.6	38.0	570	17	B15851
C 14	285.6	38.0	759	17	A0747880
C 15	285.4	38.0	415	9	AL144166
C 16	279.8	37.3	629	17	A0936381
C 17	279.6	37.2	648	17	A0115070
C 18	278.8	37.1	814	9	AU120614
C 19	278.6	37.1	853	17	A0750123
C 20	278.2	37.0	839	17	A0741664
C 21	277.4	36.9	690	14	BM978279
C 22	275.8	36.7	606	13	BM666042
C 23	273.6	36.4	586	17	A0529486
C 24	273.4	36.4	697	10	AW969594
C 25	271.8	36.2	679	10	AW969589
C 26	270	36.0	705	9	A1734305
C 27	268.6	35.8	543	10	AW006597
C 28	268.6	35.8	606	17	A0079766
C 29	268.6	35.8	664	10	AW816310
C 30	268	35.7	672	17	A0238783
C 31	266.6	35.5	459	9	A1334330
C 32	266	35.4	496	9	A1221026
C 33	263.4	35.1	679	10	AW974628
C 34	262	34.9	667	17	A0424876
C 35	262	34.9	733	17	A0424876
C 36	261.8	34.9	522	10	BE673182
C 37	261.4	34.8	535	9	A1819373
C 38	260.2	34.6	379	10	BE246065
C 39	259.6	34.6	395	10	AW629172
C 40	259.2	34.5	691	17	A0480677
C 41	258.2	34.4	510	17	A0636152
C 42	258	34.4	552	17	A0634302
C 43	256.4	34.1	597	9	AL708597
C 44	256	34.1	696	10	AW969593
C 45	254.8	33.9	763	17	AZ517838

ALIGNMENTS

RESULT 1	BM691815	515 bp	mRNA	linear	EST 28-FEB-2002
LOCUS	UI-E-CL1-aet-d-02-0-UI.r1	UI-E-CL1	Homo sapiens	CDNA clone	
DEFINITION	UI-E-CL1-aet-d-02-0-UI 5', mRNA sequence.				
ACCESSION	BM691815				
VERSION	BM691815.1	GI:19005073			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 515)				
TITLE	Bonardo,M.F., Lennon,G. and Soares,M.B.				
JOURNAL	Normalization and subtraction: two approaches to facilitate gene discovery				
MEDLINE	Genome Res. 6 (9), 791-806 (1996)				
COMMENT	9704477				
	Contact: Soares, MB				
	Program for Rat Gene Discovery and Mapping				
	University of Iowa				
	451 Eckstein Medical Research Building Iowa City, IA 52242, USA				
	Tel: 319 335 8250				
	Fax: 319 335 9565				
	Email: msoares@blue.weeg.uiowa.edu				
	Tissue Procurement: Dr. Gregg Hageman				
	CDNA Library preparation: Dr. M. Bento Soares, University of Iowa				
	CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa				
	DNA Sequencing by: Dr. M. Bento Soares, University of Iowa				





```
QY 171 AGAAGAAATTTACAGTGTACATAGAGTTAGTTGATGAAGCAGTAGCAGGATCGAGAGG 230
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 368 ACAAGGATTTAGATATTACATAAATAGTTAGTTGATGAAGCAGCAGCAGGTTTGAGAGG 309
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 231 ATCGATTCGAATTTCAAAATAGTTCTCTGTGGGTAAATGCTATCAAAATGGCGTCGCA 290
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 308 ATTGACTCCATTTTGAAGAAGCTCTACTGTGGGCAAAATGCTATCAAAACAGCATCGCA 249
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 291 TGCTACAGAGAAATCTATCATGAAGAGAGAGTCAATTCATGTGGCAAACTTCATTTGTTG 350
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 248 TGCTACAGAGAAATCTTTTCATGAAGAGAGAGTCAACTCATGCGAAAGCTTCACCGATG 189
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 351 TCGTATTTTGAAGAATTCAGGAGCAGCCCAACCTTCAACAACTGATCCCTGATCAGT 410
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 188 TCTTATTTTGAAGAATTCGCCAGCAGCCCAACCTTCAACAACTGATCCCTGATCAGT 129
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 411 CAGGAGCATCCCATTCAGGCGAGAACTCCAGCAGTAAAGATATGATTTCTTAA 470
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 128 CAGCAACATCAACTCGAGGCAAGACCTCCAGCAGCAAAAGAGTACGACTCACCGAA 69
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 471 GGATCAGATGAACATTAGCATTTTTTAAAGCAATAAAGTATTTT 516
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 68 GGCTCAGATTGATATCAGCATTTTT-AGCAATAAAGTATTTT 24
```

```
RESULT 3
N24942/c
LOCUS
DEFINITION
YX73405.s1 Soares melanocyte 2NBHM Homo sapiens cdna clone
IMAGE:267369 3', mRNA sequence.
N24942
VERSION
N24942.1 GI:1139092
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 465)
AUTHORS
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
,R., Williamson,A., Wohldmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 334
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 2431 Std Error: 0.00
Seq primer: m13 -40 forward
High quality sequence stop: 334.
Location/Qualifiers
1..465
/organism="Homo sapiens"
/db_xref="GDB:387701"
/db_xref="taxon:9606"
/clone="IMAGE:267369"
/clone_lib="Soares melanocyte 2NBHM"
/sex="Male"
/tissue.type="melanocyte"
/lab_host="DH10B (ampicillin resistant)"
note="Vector: pY73D (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cdna
was primed with a Not I - oligo(dT) primer [5,
TGTACCAATCTGAAGTGGGAGCGGCCGAGTTTATTTTATTTT 3'],
double-stranded cdna was size selected, ligated to Eco RI
```

```
BASE COUNT 117 a 93 c 95 g 160 t
ORIGIN
```

```
Query Match 41.1%; Score 309; DB 14; Length 465;
Best Local Similarity 82.2%; Pred. No. 1.6e-64;
Matches 379; Conservative 0; Mismatches 80; Indels 2; Gaps 2;
QY 57 TCATGATAAACTTTCAGTGGATGAGGAGCTGCCCTCTTAT-GATGAACAAAAGAGGTGGTT 115
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 465 TCATGATAAACTTTACAGATCAGGAGTTGTTCTTATGGATGCGCAAAAGAGGTGGTT 406
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 116 TCTTGAATGAATCTACTCTCTGGTGAAGATGCTGTGAACATTTTGAATGACAAAGAA 175
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 405 TCTTGAGATGAATCTACACCTGTGTGAAGATGCTGTGAACATTTTGAATGACAAAGAA 346
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 176 GAATTTACAGTGTACATAGATTTAGTTGATGAAGCAGTAGCAGGATTCGAGAGATCGA 235
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 345 GGATTTAGAATATTACATAAACTTAGTTGATAAAGCAGCAGGAGTTTGAGAGGATTGA 286
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 236 TTCCAATTTCAAAATAAGTTCTCTCTGGGTAAATGCTATCAATGGCGTCGATGCTA 295
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 285 CTCCAATTTTGAAGAAGAGCTCTACTGTGGCAAAATGCTATCAAAACAGCATCGATGCTA 226
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 296 CAGAGAAATCTATCATGAAGAGAGAGTCAATTTGATGTGGCAAACTTCATTTGTTGCTGTA 355
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 225 CAGAGAAATCTTTCATGAAGAGAGAGTCAACTCATGCGAAAGCTTCACCGATGCTGTA 166
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 356 TTTTGAAGAAATTTGTCAGGACCCCAACCTTCAACAACTTCACCTGATCAGTCAGGA 415
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 165 TTTTAGGAAATGGCCACAGCCCAACCTTCAGCAGCCCAACCTTCAGCAGTCAGTCAGGA 106
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 416 GCCATCCACATGAGCGGAGAGACCTCCAGCAGTAAAGATTTATGATTTCTTAAAGGATC 475
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 105 ACCATCAACCTCGAGGAGAGAGCTCCAGCAGCAAAAGAGTACGACTCACCGAGGCTC 46
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 476 AGATGAACATGAGCATTTTTTTAAAGCAATAAAGTATTTT 516
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 45 AGATGATAATCAGCATTTTTT-AGCAATAAAGTATTTT 6
```

```
RESULT 4
AZ519098
LOCUS
DEFINITION
RPCI-11-79K17.TJD RPCI-11 Homo sapiens genomic clone RPCI-11-79K17,
DNA sequence.
ACCESSION
AZ519098
VERSION
AZ519098.1 GI:10830264
KEYWORDS
GSS.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 743)
AUTHORS
Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
,J.C.
BAC end sequences of library RPCI-11
Unpublished (1997)
TITLE
Other_GSSs: RPCI11-79K17.TJD RPCI11-79K17.TV
JOURNAL
COMMENT
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieterdejong.med.buffalo.edu). Clones may be purchased from
```

FEATURES  
source

ORGANISM	Homio sapiens
Eukaryota; Metazoa;	Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria;	Primates;
	Catarrhini; Hominoidea; Homo.

RESULT 6  
VA608733/C

RESULT 6  
VA608733/C

LOCUS AA608733 460 bp mRNA linear EST 09-MAR-1998  
DEFINITION ae56f06.s1 Stratagene lung carcinoma 937218 Homo sapiens cDNA clone  
IMAGE:950915.3, similar to TR:G1224066 G1224066 ORF2: FUNCTION  
UNKNOWN. ;, mRNA sequence.  
AA608733  
ACCESSION AA608733.1 GI:2457161  
VERSION  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 460)  
REFERENCE Hillier, L., Aken, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,  
AUTHORS Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin  
J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,  
White, Y., Wylie, T., Waterston, R. and Wilson, R.  
TITLE WashU-NCI human EST Project  
JOURNAL Unpublished (1997)  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 558 Std Error: 0.00  
Seq primer: -40ml3 fwd: ET from Amersham.  
FEATURES  
source  
1. 460  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:950915"  
/clone\_lib="Stratagene lung carcinoma 937218"  
/tissue\_type="lung carcinoma"  
/cell\_line="NCI-H69"  
/dev\_stage="cell line NCI-H69"  
/lab\_host="SOIR (kanamycin resistant)"  
/note="Organ: lung; Vector: pBluescript SK-; Site\_1: EcoRI  
; Site\_2: XhoI; Cloned unidirectionally. Primer: Oligo  
dT. Small cell carcinoma cell line NCI-H69. Average  
insert size: 1.0 kb; UNIZAP XR Vector; -5' adaptor  
sequence: 5' GAATTCGGCAGGAG 3' -3' adaptor sequence: 5'  
CTCGACTTTTCTTTTCTTTT 3'."  
BASE COUNT 114 a 92 c 94 g 160 t  
ORIGIN  
Query Match 40.1%; Score 301; DB 9; Length 460;  
Best Local Similarity 81.9%; Pred. No. 1.4e-62;  
Matches 371; Conservative 0; Mismatches 80; Indels 2; Gaps 2;  
QY 65 AAACCTCAGTGGATGAGAGTGGCTCTTAT-GATGAACAAGAAGGTGTTTCTTGAAA 123  
DB 460 AAACCTTAAACAGATGAGGAGTGTCTTCTATGGATGCGCAAGAAGTGTCTTCTGAGA 401  
QY 124 TGAATCTACTCCCTGGTGAAGATGCTGTGACATTTGTAATGACACAAGAATTTAC 183  
DB 400 TGAATCTACACCTGGTGAAGATGCTGTGACATTTGTAATGACACAAGAAGTATTAG 341  
QY 184 AGTGTTCATAGAGTTAGTTGATGAAGCAGTAGCAGGATTCGAGAGATTCGATTCCAATT 243  
DB 340 AATATTACATAAATCTAGTTGATGAAGCAGCAGCAGGAGGTTTGAGAGGATTCAGTCCCAATT 281  
QY 244 TCAAAATAAGTTCTTCTGTGGGTAAATGCTATCAAAATGGCGTCGATCTACAGAGAA 303  
DB 280 TTGAAAGAAGCTCTACTGTGGGCAAAATGCTATCAACAGCATCGCATCTACAGAGAA 221  
QY 304 TCTATCATGAAGAAGAGTCAATTGATGTGCGAAGCTTCATTTGTCGTATTTTAAGA 363  
DB 220 TCTTTCATGAAGAAGAGTCAACTCATCGAAGAAGCTTCACCGATGCTTATTTTAAGA 161  
QY 364 AATTGTGAGGACCAACCCCAACCTTCAACAACCATGACCTGATCAGTCAGGAGCCATCCA 423

Db 160 AATTGGCCACAGCCACCCCAACCTTCAGCAGCCACCAACCTGACCAAGTCACCAACCATCAA 101  
QY 424 CATTGAGGCGAGACCTCCAGCAGCTAAAGAGTATTGATCTCTTAAGAGTCAGATGAAC 483  
DB 100 CCTCAGGCGAAGACCTCCAGCAGCAAAAGAGTACGACTACCCGAAAGCTCAGATTGAT 41  
QY 484 ATTAGCATTTTAAAGCAATAAAGTATTTTAA 516  
DB 40 AATCAGCATTTT-AGCAATAAAGTATTTTAA 9  
RESULT 7  
LOCUS AA750429 857 bp DNA linear GSS 19-JUL-1999  
DEFINITION HS-5574\_B1\_E07\_T7A RPOI-11 Human Male BAC library Homo sapiens  
genomic clone Plate-1150 Col-13 Row-J, DNA sequence.  
AA750429  
ACCESSION AA750429  
VERSION  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 857)  
REFERENCE Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,  
AUTHORS Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and  
Hood, L.  
TITLE Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
MEDLINE 99380589  
COMMENT Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Clones are derived from the human BAC library RPOI-11. For BAC  
library availability, please contact Pieter de Jong  
pieter@delong.med.buffalo.edu. Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering\_bac.htm)  
or from Research h Genetics (info@resgen.com). BAC end Web Server:  
http://www.htsc.washington.edu  
plate: 1150 row: J column: 13  
Seq primer: T7  
Class: BAC ends  
High quality sequence stop: 857.  
FEATURES  
source  
Location/Qualifiers  
1. 857  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="Plate-1150 Col-13 Row-J"  
/clone\_lib="RPOI-11 Human Male BAC Library"  
/sex="male"  
/note="Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
Male blood DNA was isolated from one randomly chosen donor  
and partially digested with a combination of EcoRI and  
EcoRI Methylase. Size selected DNA was cloned into the  
pBACe3.6 vector at EcoRI sites"  
BASE COUNT 273 a 156 c 172 g 255 t 1 others  
ORIGIN  
Query Match 39.5%; Score 297; DB 17; Length 857;  
Best Local Similarity 76.9%; Pred. No. 1.4e-61;  
Matches 442; Conservative 0; Mismatches 120; Indels 13; Gaps 6;  
QY 57 TCATGATAAACTTCAGTGGATGAGGAGTGGCTCTTATCA-TGAACAAGAAGTGGTT 115  
DB 108 TCATGATCAAACTTCAGTGGATGAGGAGTGGCTCTTGTGTAATGAGCAAGAAGTGGTT 167  
QY 116 TCTTGAATGGAATCTACTCCTGGTGAAGATGCTGTGAACATGTTTGAATGACAAGAA 175

```
Db 168 TCTTGAGATGAATCTACTCTCGTGAAGATGCTGTGTACACTGTTGAAATGACAAGAA 227
QY 176 GAATTTACAGTTTACATAGAGTTAGTTGATCAACAGCAGTAGCAGGATTCGAGAGGATCGA 235
Db 228 GGAATTTATTTATACATATCTTCTAGTTGATAAATCAGCAGCAGGCTTTGAGAGGATTCGA 287
QY 236 TTCCAATTTCAAAATAGTCTTCTGTGGTAAATGCTATCAAAATGGCGTCCATGCTA 295
Db 288 CTCCAATTTGAAAGAGTTCTGCTATGTTAGAGGGCTATCAACAACATTTGTAIGCTA 347
QY 296 CAGAAATCTATCATGTAAGAGAGAGTCAATTTGATGTGGCAAACTTCATTTGTGCTGA 355
Db 348 CAGAGAAATCTTTCTGTGA---AAGAGTCAATCAATGAGGCAAACTTTCATTTGTGCTTA 403
QY 356 TTTTAAAGAAATGTCAGGACCAACCCACCTTCAACACCATGACCCCTGATCAGTCAGGA 415
Db 404 TATTAGAAATTTGCTCAGCCACCCCAACTTTTGGCAACCCACCA-CTGATCAGTCAGCA 462
QY 416 GCCATCCATTTGAGGCGAGAACCTCCAGCAGTAATAAGATTATGATCTCTTAAGAGATC 475
Db 463 GCCATCAACATCAGGCGACACCCCTCCATTTAGCAAAAG---AGCTTGTCTGAAGGGTC 519
QY 476 AGATGAACATTAGCATTTTTTTTAAAGCAATAAAGTATTTTACGTAAAGATATGATGTTAT 535
Db 520 AGATGATCATTTAGCATTTTCTCAGCAATAAAGCATTTTAAATTAAGCCATGTA---CAT 576
QY 536 TTTTGGCAATAATGCTATTATGCTATTTATGACTTCCAGTATATGTTAAACATAACTTT 595
Db 577 TTTTGGCAATAATGCTATTATGCTATTTTAAATTAAGCATTTTAAATTAAGCCATGTA---CAT 636
QY 596 AA-ATGACATGGGAGATAAAGTATTTTGCCTTTT 629
Db 637 AACATGATTTGGGAAACCAAAAATTTGTACATTT 671
```

```
RESULT 8
LOCUS AQ312301 673 bp DNA linear GSS 04-MAY-1999
DEFINITION RPCI11-103G24.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-103G24,
DNA sequence.
```

```
ACCESSION AQ312301
VERSION AQ312301.1 GI:4043965
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 673)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P., and Venter,J.C.
TITLE Use of human BAC End Sequences for Sequence-Ready Map Building
JOURNAL Unpublished (1998)
COMMENT Other GSSs: RPCI11-103G24.TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
```

```
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pletter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html
Seq primer: Sp6
Class: BAC ends.
Location/Qualifiers
1. 673
/organism="Homo sapiens"
/db_xref="GDB:753935"
/db_xref="taxon:9606"
```

```
FEATURES
Source
```

```
Query Match 39.2%; Score 294.2; DB 17; Length 673;
Best Local Similarity 70.5%; Pred. NO. 6.6e-61;
Matches 470; Conservative 0; Mismatches 178; Indels 19; Gaps 5;
```

```
QY 52 CGACTTCTATGATAAACTTCACTGGATGAGGAGCTGCTCTTATG-ATGAAACAAGAAG 110
Db 8 CCATCTCTTAGTAAACATTGAATGGATGAGGAGTCCCTCTCATGAATAAGCAAGAAG 67
QY 111 TGGTTTCTTGAATGGAATCTACTCTCTGGTGAAGATGCTGTGAACATTTGTTGAATGACA 170
Db 68 TGGTTTCTTGAATGGAATCTACTCTCTGGTGAAGATGCTGTGAACATTTGTTGAATGACA 127
QY 171 AGAAGAAATTTACAGTGTTCATAGATTTAGTTGATGAAGCAGTACGAGGATTCGAGG 230
Db 128 ACAAGGATTTAGAACATTTCCATAAACTTCGTTGATAAAAGTAGTGGCAGGATTTGAGACA 187
QY 231 ATCGATTCCAAATTTCAAAATAGTTCTCTCTGGTGAAGATGCTATCAAAATGCGTTCGCA 290
Db 188 CTGAATCTCCAACTTTGAAGAAGATTTCTACTTTGGGTAAATGTTATCAAAATAGCATACA 247
QY 291 TCGTACAGAGAAATCTATCATGAAAGGAGAGTCAATTTGATGGCAAACTTCATCTTCTG 350
Db 248 TGCTACAGAGAAACCTTTTGCAAAAGG-AGACCAATCAATGCAGCAAGTTTCATTCCTG 306
QY 351 TCGTATTTTAAAGAAATTTCTCAGGACCCCAACCTTCAACACCATGACCCCTGATCAGT 410
Db 307 TCTTATTTTAAAGAAATTAACACAGCCACTCCAACTTCCAGCAACCAAAATTTTGGTCAGT 366
QY 411 CA-----GGAGCCATCCCATTCAGGCGGAGAACTCCAGCAGTAAAAGATTATGATT 463
Db 367 CAGTAGTCCGTAGCCATCCACAGGAGTAAGACCTCCACCTGAAATTAATATGACT 426
QY 464 CTCTAAAGGATCAGATGAACATTTAGCATTTTAAAGCAATAAAGTATTTTTCAGTAAGA 523
Db 427 CACTAAGGTTCAAGTAATTTAGCATTTTCAATATGCTATTTTAAATCAAAAGTG 486
QY 524 TATGATATGTTATTTTATAGGCATATGCTATTTATGATTTATAGACTCCAGTATATGTT 583
Db 487 TGTACAT-TGTCTTTTATAGACATATGTTGACACCTTAATACACCAAGTATAGTGT 545
QY 584 AAACATTAACTTTAAATGCACCTGGGAGATAAAGTATTTG-----CTCTTTTATGAT 634
Db 546 AAACCTGACTTTTATATGCACTGGGCAACCAAAATTTGTGTGACTCATTTTATATGAC 605
QY 635 ATTGTCTTTATTCAGTAGTCTGTAATGGAACACTACATATCTCTTGGTACACCTGTAT 694
Db 606 ATTGGCTTTTATTTAGTGTAGTCTGGAACCTTAACCTGCAGTATCTCTAAAGTATGCTGTAT 665
QY 695 ACAGAAA 701
Db 666 TAGAAA 672
```

```
RESULT 9
LOCUS AL521707/c 914 bp mRNA linear EST 13-FEB-2001
DEFINITION AL521707 LRI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB003YF15 3
prime, mRNA sequence.
ACCESSION AL521707
VERSION AL521707.1 GI:12785200
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

REFERENCE 1 (bases 1 to 914)  
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
Bp 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
Location/Qualifiers  
1..914  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0B003YF15"  
/clone\_lib="LTI\_NFL004\_NBC2"  
/sex="male"  
/tissue\_type="neuroblastoma cells"  
/lab\_host="DH10b"  
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 234 a 197 c 178 g 300 t 5 others  
ORIGIN

Query Match 38.7%; Score 290.6; DB 9; Length 914;  
Best Local Similarity 82.0%; Pred. No. 5.2e-60;  
Matches 346; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

Qy 52 CGACTTCATGATAAACTTCAGTGGATGAGGAGCTGCCTCTTAT-GATGAACAAGAGG 110  
Db 423 CAATCTCATGATAAACTTTAACAGATGAGGAGTTGTTCTTATGGATGCGCAAGAAG 364

Qy 111 TGGTTCTTGAATGAATCTACTCTGTGGAAGATGCTGTGAACATTTGTAATGACA 170  
Db 363 TGGTTCTTGAATGAATCTACTCTGTGGAAGATGCTGTGAACATTTGTAATGACA 304

Qy 171 AGAAGAATTTACAGTTTACATAGAGTTAGTTGATGAAGCAGTGGAGGATTCGAGAG 230  
Db 303 ACAAGAGATTTAGATATATACATAAACTTAGTTGATGAAGCAGGAGTTGAGAGG 244

Qy 231 ATCGATTCCAATTTCAAAATAGTTCTCTGTGGGTAATGCTATCAAAATGGCGTCGA 290  
Db 243 ATTGACTCCAATTTGAAAGAGAGCTCTACTGTGGGCAAAATGCTATCAAAATGGCG 184

Qy 291 TGCTACAGAGAAATCTATCATGAAGAGAGTCAATGATGTGGCAAACTTCATTTGTTG 350  
Db 183 TGCTACAGAGAAATCTTTCATGAAGAGAGTCAATGATGTGGCAAACTTCATTTGTTG 124

Qy 351 TCGTATTTTGAAGAAATTCAGGAGCAACCCCAACCTTCAACACCATGACCTGATCAGT 410  
Db 123 TCTTATTTTAGAATTTGCCAGAGCAACCCCAACCTTCAACACCATGACCTGATCAGT 64

Qy 411 CAGAGGATTCATGAGGAGGAGAACTCCAGCAGTAAAGATTTATGATTTCTTAA 470  
Db 63 CAGAGGATTCATGAGGAGGAGAACTCCAGCAGTAAAGATTTATGATTTCTTAA 4

Qy 471 GG 472  
Db 3 AG 2

RESULT 10  
BM703722  
LOCUS  
DEFINITION UI-E-CL1-aff-h-03-0-UI.r1 UI-E-CL1 Homo sapiens cDNA clone  
UI-E-CL1-aff-h-03-0-UI 5', mRNA sequence.

BM703722  
BM703722.1 GI:19016980  
EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 480)  
AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
COMMENT Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msoares@blue.weeg.uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).  
The following repetitive elements were found in this cDNA sequence: 1-480, >TIGER14DNA/MER2\_type  
Seq primer: M13 Reverse.

FEATURES  
Location/Qualifiers  
1..480  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="UI-E-CL1-aff-h-03-0-UI"  
/clone\_lib="UI-E-CL1"  
/tissue\_type="human retina"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)."  
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-CL1 is a normalized cDNA library containing the following tissue(s): retina. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CCGCG. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

BASE COUNT 169 a 98 c 103 g 110 t  
ORIGIN

Query Match 38.7%; Score 290.4; DB 14; Length 480;  
Best Local Similarity 82.7%; Pred. No. 5.1e-60;  
Matches 344; Conservative 0; Mismatches 71; Indels 1; Gaps 1;

Qy 52 CGACTTCATGATAAACTTCAGTGGATGAGGAGCTGCCTCTTAT-GATGAACAAGAGG 110  
Db 62 CAATCTCATGATAAACTTTAACAGATGAGGAGTTGTTCTTATGGATGCGCAAGAAG 121

Qy 111 TGGTTCTTGAATGAATCTACTCTGTGGAAGATGCTGTGAACATTTGTAATGACA 170  
Db 122 TGGTTCTTGAATGAATCTACTCTGTGGAAGATGCTGTGAACATTTGTAATGACA 181

Qy 171 AGAAGAATTTTACAGTTTACATAGAGTTAGTTGATGAAGCAGTGGAGGATTCGAGAG 230  
Db 182 ACAAGAGATTTAGATATATTTACATAAACTTAGTTGATAAAGCAGCAGGAGTTGAGAG 241

Qy 231 ATCGATTCCAATTTCAAAATAGTTCTTCTGTGGTAAATGCTATCATCAATGGCGTCGA 290

```

Db 242 ATTGACTCCAAATTTGAAGAAGCTCTACTGTGGCAAAATGCTATCAACAGCATCGCA 301
QY 291 TGTACAGAGAAATCTATCATGAAGAAGAGTCAATGTGTGGCAAACTTCATTTGTG 350
Db 302 TGTACAGAGAAATCTTCTATGAAGAAGAGTCAATGTGTGGCAAACTTCACCGATG 361
QY 351 TGTATTTTAAAGAAATGTGTCAGGACCAACCCCAACCTTCAACAACCATGACCTGATCAGT 410
Db 362 TCTATTTTAGGAAATGTCACAGCCACCCCAACCTTTCAGAGCCACCCACCTTGACCACT 421
QY 411 CAGGAGCCATCCACATTTAGGCGGAGAGACCTCCACAGCTATAAAGATTTATGATTC 466
Db 422 CAGCAACCATCAACCTCGAGCGAAGACCTCCACACGACCAAAAGAGTACGACTCAC 477

RESULT 11
A1189864/c
LOCUS
DEFINITION
q23h08.x1 Soares_placenta_8to9weeks_2NBP8to9W Homo sapiens cDNA
clone IMAGE:1724607 3' similar to TR:Q13538 Q13538 ORF2: FUNCTION
UNKNOWN. ; mRNA sequence.
ACCESSION
A1189864
VERSION
A1189864.1 GI:3741073
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 512)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ccgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: ccgaps-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (infoimage.llnl.gov) for further information.
Insert Length: 1099 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 427.
FEATURES
Source
Location/Qualifiers
1..512
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1724607"
/dev_stage="two placenta: one from 8 weeks and another
from 9 weeks post conception"
/lab_host="DH10B (ampicillin resistant)"
/notes="Organ: placenta; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAGTGGGAGCGCCGCGGATTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo."
BASE COUNT 145 a 87 c 92 g 188 t
ORIGIN
Query Match 38.6%; Score 289.8; DB 9: Length 512;
Best Local Similarity 76.8%; Pred. No. 7.3e-60;
Matches 394; Conservative 0; Mismatches 112; Indels 7; Gaps 3;

QY 103 AAGAAGTGGTCTTCTGAATGAATCTACTCTGTTGAAGATGCTGTGAACATGTTG 162
Db 511 AACCAAGTGGTCTTCTGAGATGAATCTACTCTGTTGAAGATGCTGTGAACATGTTG 452
QY 163 AAATGACAAAGAAATTTACAGTGTACATAGAGTGTATGATGAACAGCAGTACGAGAT 222
Db 451 AAATGACAAAGAAATTTACAGATTTACATATAATTTAGTTGATAAAGCAGTGGCAGTT 392

```

```

QY 223 TCGAGAGATCGATTCCAAATTTCAAAATAAGTTCTCTCTGGTAAATGCTATCAATG 282
Db 391 TTGAGAAGATAGCTCCAGTTTGAAGAAGAGTTCTACTGTGGTAAATGCTATCAACA 332
QY 283 GCGTCGATGTCAGAGAGAAATCTATCAATGAAGAAGAGTCAATTTGATGTGGCAACTT 342
Db 331 GCATCACACATTCAGATATAATCTTTTGTGAAAGAAGAAAGTCAATTTGTTCACAAC-T 273
QY 343 CATTTGTTGCTGATTTTAAAGAAATTTGTCAGGACCCACCCCAACCTTCAACAACCATGACCC 402
Db 272 CTTCAATTTGCTATATTTTAAAGAAGTTGCCACAGCCACCCCAACCTTTCAGCAACCATGACCC 213
QY 403 TGATCACTCAGGAGCCATCCACATTTGAGCGGAGAGACCTCCACAGCAGTAAAGATTTATGAT 462
Db 212 AGATCAATCAGCAGCCATCAATATCAAGGCTAGCCCTCCACAGCAAAAGACCATC 153
QY 463 TCTCTAAAGGATCAGATGAACATTAGCATTTTAAAGCAATAAAGTATTTTACCTAAG 522
Db 152 TTCGAGAAGCTCAGATGAACATTAGCATTTTAAAGCAATAAAGTATTTTGAATGAG 93
QY 523 ATATGATGTTATTTTAAAGCAATAAAGTATTTTAAAGCAATAAAGTATTTTGAATGAG 582
Db 92 ATATGTA---TATTTTCAAGCCATTAATGTTTACACACTTACTAGATTACAATAATA 36
QY 583 TAAACATAACTTTTAAATGCACTGGGAGATAAAA 615
Db 35 TAAATGCT---TTTATATGCTCGCTGGGAAAAAAA 6

```

```

RESULT 12
A0237336
LOCUS
DEFINITION
A0237336
ACCESSION
A0237336
VERSION
A0237336.1 GI:3669627
KEYWORDS
GSS.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 542)
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
Use of human BAC end Sequences for Sequence-Ready Map Building
Unpublished (1998)
Other GSSs: RPC111-66L13.TK
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library RPC11-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1..542
/organism="Homo sapiens"
/db_xref="GDB:7525236"
/db_xref="taxon:9606"
/clone="RPC11-11-66L13"
/clone_lib="RPC11-11"
/sex="Male"
/cell_type="Lymphocytes"
/notes="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;

```







QY 406 TCAGTCAGAGCAGCATCCAGCATTTAGGCGGAGAACTCCAGCAGTAAAAAGATTATTGATTCT 465  
Db 115 CCAGTCAGAGCACTCAAGCTCGAGCGAAGACCTCCACGAGAAAAGAGTAGACTCA 56  
QY 456 CTAAAGGATCAGATGAACATTAGCATTTTTTTTAAAGCAATAAAGATTATTTTA 516  
Db 55 CCGAAGGCTCAGATTGATTAATCAGCATTTTTTT-AGCAATAAAGATTATTTTA 6

Search completed: April 11, 2003, 22:37:47  
Job time : 844.08 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 15:37:18 ; Search time 20.7377 Seconds  
(without alignments)  
11106.080 Million cell updates/sec

Title: US-09-914-152-3\_COPY\_2000\_2750

Perfect score: 751

Sequence: 1 cctgccttggtcctccaaag.....tcatggagcgtgaaagatgc 751

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.\*

1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*

2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*

3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*

4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*

5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*

6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	239.8	31.9	112132	4	US-09-741-150-3
2	232.8	31.0	80246	4	US-09-078-294-4
3	224	29.8	80595	4	US-09-078-294-3
4	188.6	25.1	14753	4	US-09-821-736-3
5	137.2	18.3	65042	4	US-09-784-316-3
6	96.2	12.8	533	4	US-09-105-542A-11
7	92.4	12.3	313	4	US-08-991-789A-234
8	92.4	12.3	313	4	US-09-062-451-234
9	92.4	12.3	313	4	US-09-598-326-234
10	90.4	12.0	229	4	US-08-905-223-138
11	83.6	11.1	3646	4	US-08-991-789A-227
12	83.6	11.1	3646	4	US-09-062-451-227
13	83.6	11.1	3646	4	US-09-598-326-227
14	60.8	8.1	4396	4	US-09-821-736-1
15	54.2	7.2	1001	4	US-09-641-638-144
16	50.4	6.7	19011	1	US-08-310-356-36
17	50.4	6.7	19557	5	PCT-US92-06300-1
18	47.8	6.4	21234	4	US-08-810-671-3
19	47	6.3	5543	2	US-08-687-080-101
20	46.4	6.2	1442	2	US-08-454-557C-120
21	46.4	6.2	1442	2	US-08-340-426D-120
22	46.4	6.2	1442	2	US-08-450-673C-120
23	46.4	6.2	19736	4	US-09-740-035-3
24	46.4	6.2	112132	4	US-09-741-150-3
25	46.2	6.2	3609	4	US-09-705-299-11
26	46.2	6.2	87350	3	US-08-781-891-79
27	46.2	6.2	87543	4	US-09-791-211-3

28 45.8 6.1 2336 1 US-08-247-946A-1 Sequence 1, Appli  
29 45.8 6.1 2336 5 PCT-US95-06420-1 Sequence 1, Appli  
30 45.8 6.1 5359 2 US-09-070-060-1 Sequence 1, Appli  
31 45.8 6.1 5372 3 US-09-357-746-1 Sequence 1, Appli  
32 45.8 6.1 70000 4 US-09-851-896-3 Sequence 13, Appli  
33 45.6 6.1 19124 2 US-08-487-826B-13 Sequence 1, Appli  
34 45.4 6.0 694 1 US-08-250-314-1 Sequence 1, Appli  
35 45.4 6.0 694 1 US-08-708-107-1 Sequence 1, Appli  
36 45.2 6.0 615 4 US-09-385-982-528 Sequence 528, App  
37 45.2 6.0 1749 4 US-09-149-476-54 Sequence 54, Appl  
38 44.8 6.0 12394 4 US-09-488-856A-10 Sequence 10, Appl  
39 44.8 6.0 18596 4 US-09-318-448-11 Sequence 11, Appl  
40 44.8 6.0 84495 4 US-09-797-906-3 Sequence 3, Appli  
41 44.6 5.9 87350 3 US-08-781-891-79 Sequence 79, Appl  
42 44.6 5.9 87543 4 US-09-791-211-3 Sequence 3, Appli  
43 44.4 5.9 12047 2 US-09-022-461-1 Sequence 1, Appli  
44 44.4 5.9 12047 4 US-09-033-556-3 Sequence 3, Appli  
45 44.4 5.9 72604 4 US-09-268-992-7 Sequence 7, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-741-150-3

; Sequence 3, Application US/09741150

; Patent No. 6436689

; GENERAL INFORMATION:

; APPLICANT: GUEGLER, Karl et al

; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND

; TITLE OF INVENTION: USES THEREOF

; FILE REFERENCE: CL000968

; CURRENT APPLICATION NUMBER: US/09/741,150

; CURRENT FILING DATE: 2000-12-21

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 112132

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)...(112132)

; OTHER INFORMATION: n = A,T,C or G

US-09-741-150-3

Query Match

Best Local Similarity 31.9%; Score 239.8; DB 4; Length 112132;

Matches 484; Conservative 0; Mismatches 162; Indels 57; Gaps 9;

QY 52 CGACTTCATGATAAACTTCAGTGGATGAGGAGCTGCTCTTA-TGATGACAAAG-AAG 109  
Db 42311 CAATCTCATGTGTAACCTTGAACAGATGAGAGTGTCTTCTACAGATGACGAAGAAG 42370  
QY 110 GTGGTTCTTGAATGGAATCTACCTCGTGGTGGAGATGCTGTGAACATTTGTAATGAC 169  
Db 42371 CGGTTCTTGAATGGAATCTAGTCTCGTGGAGGATGCTATGAACCTTGTGTAATGAC 42430  
QY 170 AAGAAGA-----ATTACAGTGTACATAGAGT 198  
Db 42431 AACCTTGATGTTGTGAACCTTGTGNAATCTAAACAGATTTAGATAATTTACATAACA 42490  
QY 199 TAGTGTATGAA-GCAGTAGCAGGATTCGAGGATCGATTCCAAATTTCAAAATAGTTCT 257  
Db 42491 TAGTGTATGAAAGGACCAACAGGGTTTGAAGGATTTGACTTCAATTTTGAAGAAATCT 42550  
QY 258 TCTGTGGTAAATGCTATCAATGCGGTGCGATGCTACAGAGAAATCTATCATGAAGG 317  
Db 42551 ACGTGGCAAAATGCTATCGAATAGCAATGCAATGAGGCTATAAGAAATTTGTTTCATGAAGG 42610  
QY 318 AAGAGTCAATTTGATGTCGAACACTTCATGTTGTGCTATTTTAAAGAAATGTCAGACCA 377

Db	42611	AAGAGCTCAATAGATGAAGCAAAATTTTACTGTGGCTTATTTTAAAGAAATGCCACAGCCA	42670
Qy	378	CCCCAACCTTCAACAAACCATGACCCCTGATCAGTCAGGAGCCATGCACATTCAGGCGAGAA	437
Db	42671	CCCTAACCTTTCAGAGCCACCA-CCTGATCAGTCATCAACCCATTAATATTGAGACAAGAC	42729
Qy	438	CCTCCAGCAGTAAAGAAAGATTATGATT--CTCTAAGGATCATGATGAACATTAGCATTTTT	495
Db	42730	ACTCCACCAGCAAAATGACAACAACCTAACACTGAAGACTCAGGTGATTAGCA---TTTT	42785
Qy	456	TTAAGCATAAAGTATTTTTTACGTAAGATATGTTATGTTATTTTTTAGGCATAATGCCTATT	555
Db	42786	ATACAAGAAAGTATTTGTTTAATTAAGGCATGTACATTGTTTTTTAGACATAATGCTATT	42845
Qy	556	ATGCATTTAATAGACTCCAGTATATTGT--AAACATAAAGCTTTAAAGCCACTGGAGATA	612
Db	42846	GCACACITTAATAGACTATAGTATATTGTTGAACATAACCTTTATATGCATCGGGAACA	42905
Qy	613	AAA-----GTATTTCCTTTTTATGATATTTCCTTTATTCGATGATGCTGTGA	659
Db	42906	AAAAAAAACATACATGTGACACCTCTGTTGCAAAATTTGCTTTATTTCGAGTGCTTGA	42965
Qy	660	ATGGAACACTACATTATCTCTTCCTGGGTACACCTGTATACAGAAAG	702
Db	42966	ACTGAACCCACAGTGTCTCGAGTATACCTGTATTGAGGAGG	43008

```

RESULT 2
US-09-078-294-4/c
; Sequence 4, Application US/09078294
; Patent No. 6265211
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sart, Desirée
; APPLICANT: Cancilla, Michael R.
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
; FILE REFERENCE: Davies Col
; CURRENT APPLICATION NUMBER: US/09/078,294
; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 80246
; TYPE: DNA
; ORGANISM: Nucleotide sequence of NC-contig
US-09-078-294-4

```

Query Match	31.0%;	Score 232.8;	DB 4;	Length 80246;
Best Local Similarity	67.6%;	Pred. No. 6.le-53;		
Matches	454;	Conservative	0;	Mismatches 167; Indels 51; Gaps 7;
QY	52	CGACTTCATGATAAACTTCAGTGGATGAGAGCTGCTCTTATGATGAACAAAGAAGGT	111	
Db	1337	CAATCTCATGATAAACTTGAACAGATGAGGAATTTCTTCITATGGATGACAAAGATTC	1278	
QY	112	GGTTC-----TTCAATGGAATCTACTCTCGTGAAGATCGCTGGAACATGTTGAAAT	166	
Db	1277	CATCTCAGGAATAGAGATGGAATCTACTCTGGTGAACACACTCTGAAAATTTGTTGAAAT	1218	
QY	167	GACAAGAAAGAATTTACAGTCTTACATAGAGTTAGTTGATGAACAGTAGCAGGATTCGA	226	
Db	1217	GACAACAATGGATTCAGAAATATATATTAACCTAGGTTGTTAAACAGCAGCAGGGTTGA	1158	
QY	227	GAGGATCGATTCCAATPTTCAAAATAAGCTTCTCTGTCGGTAAATGCTATCAAAATGCGGT	286	
Db	1157	GAGGATTGACTCCAAATTTCAAGAAGTCTTACTGTGCATAAATGCTATCAAAAAGCAT	1098	
QY	287	CGATGCTACAGAGAAATCTATCATGAAGGAAGAGTCAATTGATGTGGCAAACTTCATT	346	
Db	1097	CACACACTAGAGAGAAATCTTTTCATGAAGAGAGAGTCAACTGATATTGCAATCTTCACT	1038	
QY	347	GTTTGCATTTTAAAGAAATGTGAGGACCACCCCACTTTCAACAACCATGACCCCTGAT	406	

Db	1037	GTTG-----TTTTAAGAAACTGCCACAGCCACCCCTAAGCTTCAGCACCTCGACCCCC-----	987
QY	407	CAGTCAGGAGCCATCCACATTGAGCGGAGAACCTCCAGCAGTAAAAAGATTATGATTCTC	466
Db	986	-----GATCAGTCATCAATAGAAGAAAGACCTCCACCAGCAAAAAGATTACAACCTAC	934
QY	467	TAAAGGATCAGATCAACATTTAGCATTTTTTTTAAAGCAATAAAGTATTTTTTACGTAAGATAT	526
Db	933	TGAATGCTCAATGATGTTAGCATTTTTT--AGTAATAAGCATTTTTTAAAAATTA---AG	879
QY	527	GTAATGTTTTTTTAGGCATAAATGCTATTATGCACTTTAATAGACTCCAGTATATTGTAAA	586
Db	878	GTAATGTTTTTTTAGACATAACACTATTGCACACTGAACAGACTACAGTGTAGTATAAG	819
QY	587	CATAAC-TTTAAATGCACTGGGAGATAAAGTATTG-----	622
Db	818	CAAACTTTTATATGCTAGCTGGAGGCCAAAGACTTGTGTGACTCACTTTATTGTGATAT	759
QY	623	-CPTCTTTATGATATTGCTTTATTGTCAGTAGTCTGTAAATGGAAGCACTATATCTCTTG	681
Db	758	CCACTCTTCGATATTGTTATTAATGTCGTATGGTATGGAAACCAACCCACAGTATCTCAA	699
QY	682	GGTACACCTGTA	693
Db	698	GGTATGCGCTGTA	687

RESULT 3  
 US-09-078-294-3/c  
 ; Sequence 3, Application US/09078294  
 ; Patent No. 6265211  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Choo, Kong-Hong Andy  
 ; APPLICANT: Du Sart, Desiree  
 ; APPLICANT: Cancilla, Michael R.  
 ; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE  
 ; FILE REFERENCE: Davies Col  
 ; CURRENT APPLICATION NUMBER: US/09/078.294  
 ; CURRENT FILING DATE: 1998-05-13  
 ; NUMBER OF SEQ ID NOS: 29  
 ; SOFTWARE: Patentin Ver. 2.0  
 ; SEQ ID NO 3  
 ; LENGTH: 80595  
 ; TYPE: DNA  
 ; ORGANISM: Nucleotide sequence of HC-contig  
 US-09-078-294-3

Query Match	29.8%;	Score 224;	DB 4;	Length 80595;
Best Local Similarity	67.7%;	Pred. No. 1.4e-50;		
Matches 455;	Conservative 0;	Mismatches 165;	Indels 52;	Gaps 8;
QY 52	CGACTTCATGATAAACTTTCAGTGGATGAGGAGCTGCCTCTATGATGAACAAAGAGGT	111		
Db				
1626	CAATCTCATGATAAAACTTCAACAGATGAGGAATTCCTTATGATGAGCAAAAGATTC	1567		
QY 112	GTTTTC-----TTGAAATGGAATCTACTCTCHGTGGAAGATGCTGTGACATTTGTTGAAAT	166		
Db				
1566	CATCTCAGGAAATGAGATGGAATCTACTCTGTTGGAACACACTGTGAAAATTTGTTCAAAT	1507		
QY 167	GACAAGAAGAATTTACAGTGTTCATAGAGTTAGTTGATGAAGCAGCTAGCAGGATTCGA	226		
Db				
1506	GACAACAATGGATTCAGATATATTATTAAGTTAGTTGTTAAAGCAGACGAGGTTTGA	1447		
QY 227	GAGGATCGATTCCAAATTTCAAATAAGTTCTCTCTGTTGGTAAAAATGCTATCAAAATGGCGT	286		
Db				
1446	GAGGATTGA-TCCATTTTCAAAGAGATTCTACTGTGCATAAAAATGCTATCAAAAGCAT	1388		
QY 287	CGCATGCTACAGAGAATCTATCATGAAAGGAAGAGTCAAATTGATGTGGCAAACTTCATT	346		
Db				
1387	CACACACTAGAGAAAATCTTTCATGAAGGAAGAGTCAACTGATATTGCAATCTTCACT	1328		
QY 347	GTTGTGCTATTTTAGAAAATTTFCAGGACCACCCCAACCTTCAACACCATGACCTGAT	406		
Db				



```
; APPLICANT: Bullerdick, Jörn
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES OF GENES OF THE
; FILE REFERENCE: HIGH MOBILITY GROUP PROTEINS AND USES THEREFORE
; CURRENT APPLICATION NUMBER: US/09/105,542A
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: PCT/DE96/02494
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 533
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-09-105-542A-11

Query Match      12.8%; Score 96.2; DB 4; Length 533;
Best Local Similarity 80.7%; Pred. No. 6e-17;
Matches 138; Conservative 0; Mismatches 28; Indels 5; Gaps 2;

QY 358 TTAAGAAATTGTCAGGACCCACCCCAACCTTCAACAACCATGACCTGATCAGTCAGGAGC 417
Db 251 TTAAGAAATTGTCAGTCCACCCCAACCTTCAAGCAAAACCCCTGATCAATCCGCACG 310
QY 418 CATCCACATTGAGCGGAGACCTC---CAGCAGTAAAGAGATTATGATCTCTTAAAGAT 474
Db 311 CATCACACTGAGCGAGACCCCTCCCTTCAACGCAAAAGGATTACGACTCACTGAAGTT 370
QY 475 CAGATGAACATTAGCATTTCTT---AGCAATAAGTATTTTAAATAGGTA 525
Db 371 CAGATGATCATGACATTTCT---AGCAATAAGTATTTTAAATAGGTA 419

RESULT 7
US-08-991-789A-234
; Sequence 234, Application US/08991789A
; Patent No. 6225054
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; APPLICANT: Smith, John M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; NUMBER OF SEQUENCES: 292
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Ip Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,789A
; FILING DATE: 11-Dec-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 210121.419C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 234:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; SEQUENCE DESCRIPTION: SEQ ID NO: 234:
US-08-991-789A-234

Query Match      12.3%; Score 92.4; DB 4; Length 313;
Best Local Similarity 81.5%; Pred. No. 5.3e-16;
Matches 119; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

QY 57 TCATGATAAACTTCAGTGGATGAGGAGCTGCTCTTAT-GATGAACAAGAGAGTGGTT 115
Db 17 TGATGATAAACTTCAGTGGATGAGGAGCTGCTCTTATGATGAGCAAGAGAGTGGTT 76
QY 116 TCTTGAATGAATCTACTCTCTGCTGAGATGCTGTGAACATTTGTTCAATGACAAGAA 175
Db 77 TCTTGTGATGAATCTGCTCTGCAAAATGCTGTGAACGTTGTTCAAAAGACAACAA 136
QY 176 GAATTTACAGTGTACATAGAGTTAG 201
Db 137 GAGTTAGAGTAGTACATAAATTTAG 162

RESULT 8
US-09-062-451-234
; Sequence 234, Application US/09062451
; Patent No. 6344550
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; NUMBER OF SEQUENCES: 297
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/062,451
; FILING DATE: 04-APR-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.419C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 234:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-09-062-451-234

Query Match      12.3%; Score 92.4; DB 4; Length 313;
Best Local Similarity 81.5%; Pred. No. 5.3e-16;
Matches 119; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

QY 57 TCATGATAAACTTCAGTGGATGAGGAGCTGCTCTTAT-GATGAACAAGAGAGTGGTT 115
```

Db 17 TGATGATAAACTTGAATGATCAATAGTTGCTTCTTATGATGAGCAAGAAAGTAGTT 76  
Qy 116 TCTTGAATGAATGGAATCTCTCTGCTGGAAGATGCTGTGAACATTTGTGAATGACAAGAA 175  
Db 77 TCTTGTGATGAATCTCTCTGCTGGAAGATGCTGTGAACATTTGTGAATGACAAGAA 136  
Qy 176 GAATTTACAGTTTACATAGAGTTAG 201  
Db 137 GAGTTTAGAGTAGTACATAAATTAG 162

RESULT 9  
US-09-598-326-234  
; Sequence 234, Application US/09598326  
; Patent No. 6423496  
; GENERAL INFORMATION:  
; APPLICANT: Frudakis, Tony N.  
; Smith, John M.  
; Reed, Steven G.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TREATMENT AND DIAGNOSIS OF BREAST CANCER  
; NUMBER OF SEQUENCES: 247  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed Intellectual Property Law Group PLLC  
; STREET: 701 Fifth Avenue, Suite 6300  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/598,326  
; FILING DATE: 20-Jun-2000  
; CLASSIFICATION: <unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Potter, Jane E.R.  
; REGISTRATION NUMBER: 33,332  
; REFERENCE/DOCKET NUMBER: 210121.419D1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 234:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 313 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; SEQUENCE DESCRIPTION: SEQ ID NO: 234:

US-09-598-326-234  
Query Match 12.3%; Score 92.4; DB 4; Length 313;  
Best Local Similarity 81.5%; Pred. No. 5.3e-16;  
Matches 119; Conservative 0; Mismatches 26; Indels 1; Gaps 1;  
Qy 57 TCATGATAAACTTGAATGATGAGAGCTGCTCTTAT-GATGACAAGAAAGGTGGTT 115  
Db 17 TGATGATAAACTTGAATGATCAATAGTTGCTTCTTATGATGAGCAAGAAAGTAGTT 76  
Qy 116 TCTTGAATGAATCTACTCTCTGCTGGAAGATGCTGTGAACATTTGTGAATGACAAGAA 175  
Db 77 TCTTGTGATGAATCTCTCTGCTGGAAGATGCTGTGAACATTTGTGAATGACAAGAA 136  
Qy 176 GAATTTACAGTTTACATAGAGTTAG 201  
Db 137 GAGTTTAGAGTAGTACATAAATTAG 162

RESULT 10  
US-08-905-223-138/c  
; Sequence 138, Application US/08905223  
; Patent No. 6222029  
; GENERAL INFORMATION:  
; APPLICANT: Edwards, Jean-Baptiste D.  
; APPLICANT: Duclert, Aymeric  
; APPLICANT: Lacroix, Bruno  
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS  
; NUMBER OF SEQUENCES: 503  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/905,223  
; FILING DATE:  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 138:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 229 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: DOUBLE  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: CDNA  
; ORIGINAL SOURCE:  
; ORGANISM: Homo Sapiens  
; TISSUE TYPE: Brain  
; FEATURE:  
; NAME/KEY: sig\_peptide  
; LOCATION: 101..157  
; IDENTIFICATION METHOD: Von Heijne matrix  
; OTHER INFORMATION: score 4.1  
; OTHER INFORMATION: seq FVVFSTMTASSP/GE

US-08-905-223-138  
Query Match 12.0%; Score 90.4; DB 4; Length 229;  
Best Local Similarity 78.0%; Pred. No. 1.6e-15;  
Matches 103; Conservative 4; Mismatches 25; Indels 0; Gaps 0;  
Qy 127 AATCTACTCTCTGCTGGAAGATGCTGTGAACATTTGTGAATGACAAGAAAGTAGTT 186  
Db 167 AGTCTCTCTCTGCTGGAAGATGCTGTGAACATTTGTGAATGACAAGAAAGTAGTT 108  
Qy 187 GTTACATAGAGTTAGTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 246  
Db 107 ATTCATAAAGTTAGTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 48  
Qy 247 AAATAAGTTCTT 258  
Db 47 AAGAAGTTCTT 36

RESULT 11  
US-08-991-789A-227  
; Sequence 227, Application US/08991789A

```

? Patent No. 6225054
? GENERAL INFORMATION:
? APPLICANT: Fridakis, Tony N.
? Smith, John M.
? Reed, Steven G.
? TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
? TREATMENT AND DIAGNOSIS OF BREAST CANCER
? NUMBER OF SEQUENCES: 292
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Seed IP Law Group
? STREET: 701 Fifth Avenue, Suite 6300
? CITY: Seattle
? STATE: Washington
? COUNTRY: USA
? ZIP: 98104-7092
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/991,789A
? FILING DATE: 11-Dec-1997
? CLASSIFICATION: <Unknown>
? ATTORNEY/AGENT INFORMATION:
? NAME: Potter, Jane E. R.
? REGISTRATION NUMBER: 33,332
? REFERENCE/DOCKET NUMBER: 210121.419C3
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (206) 622-4900
? TELEFAX: (206) 682-6031
? INFORMATION FOR SEQ ID NO: 227:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 3646 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? SEQUENCE DESCRIPTION: SEQ ID NO: 227:
?
? US-08-991-789A-227
?
? Query Match 11.1%; Score 83.6; DB 4; Length 3646;
? Best Local Similarity 69.8%; Pred. No. 2.9e-13;
? Matches 113; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
?
? QY 353 GTATTTTAAAGAAATGTCAGGACCAACCCACCTTCAACACACCATGACCCCTGATCAGTCA 412
? Db 3459 GAATTCGTCAATTTGATCAGACGCCCTCTACCCCTTCAGCAACCAACCACCTGATCAGTCA 3518
?
? QY 413 GGAGCCATCCACATTTGAGGGGAGAACCTCCAGCAGTAAAAGATTATCATCTCTAAAGG 472
? Db 3519 GCAGCCATCAGCACCGGAGCAAGCCCTCCACCAGCAAAAAGATTCTGACTCACTGAAGA 3578
?
? QY 473 ATCATGATCACTAGTACATTTTTTAAAGCAATAAAGATTATT 514
? Db 3579 CTTGGATGATCATTAGTATTTTAAAGCAATAAAGTTTTTT 3620
?
? RESULT 12
? US-09-062-451-227
? Sequence 227, Application US/09062451
? Patent No. 6344550
? GENERAL INFORMATION:
? APPLICANT: Fridakis, Tony N.
? APPLICANT: Smith, John M.
? APPLICANT: Reed, Steven G.
? TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
? TREATMENT AND DIAGNOSIS OF BREAST CANCER
? NUMBER OF SEQUENCES: 297
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: SEED AND BERRY LLP
? STREET: 6300 Columbia Center, 701 Fifth Avenue
? CITY: Seattle
? STATE: Washington

```

```
; REFERENCE/DOCKET NUMBER: 210121.419D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 227:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3646 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 227:
US-09-598-326-227

Query Match      11.1%; Score 83.6; DB 4; Length 3646;
Best Local Similarity 69.8%; Pred. No. 2.9e-13;
Matches 113; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 353 GTATTTTAAGAAATGTCAGGACCAACCCCAACCTTCAACAACCATGACCTGATCAGTCA 412
Db 3459 GAATTCGTCGAATTCATCAGACCCCTCTACCTTCAGCAACCAACCCCTGATCAGTCA 3518

QY 413 GGAGCCATCCACATTCAGGCGAGAACCTCCAGCAGTAAAGAAAGATTATGATTCCTTAAGG 472
Db 3519 GCAGCCATCAGCAGCGAGCGAGGCCCTCCACAGCAAAAGATTCTGACTCACTGAAGA 3578

QY 473 ATCAGATGAACATTAGCATTTTAAAGCAATAAAGTATTTT 514
Db 3579 CTGTCGATCATATTAGTTTAAAGCAATAAAGTATTTT 3620

RESULT 14
US-09-821-736-1/c
; Sequence 1, Application US/09821736
; Patent No. 6326182
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
; FILE REFERENCE: CL001216
; CURRENT APPLICATION NUMBER: US/09/821.736
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4396
; TYPE: DNA
; ORGANISM: Human
US-09-821-736-1

Query Match      8.1%; Score 60.8; DB 4; Length 4396;
Best Local Similarity 58.1%; Pred. No. 4.1e-07;
Matches 129; Conservative 0; Mismatches 87; Indels 6; Gaps 1;

QY 461 ATCTCTAAGGATCAGATGAACATTTAGCATTTTAAAGCAATAAAGTATTTTACGTA 520
Db 4144 ATCATCTGAGGCTCAGATGATGTTAGCATTTTTCAGCAACAAGGTATTTTAAATTA 4085

QY 521 AGATATGATATGTTATTTTATGAGCAATGCTATTAATGATTAATGACTCCAGTATAT 580
Db 4084 GGTATGATACATGCTTTTAAAGACATATGCTATGACACCTTATAAATATATAG 4025

QY 581 TGTAAACATAACTTTAAATGCATCGGAGATAAAGTATTTGCTCTTTTATGATATTTGC 640
Db 4024 TGGAAATGTAACCTTTTACATGCCTAAGGAAAGCAAAAATTC-----ATGTCNCTGC 3971

QY 641 TTTATGTCAGTCTGTCGTAATGGAACACTACATTATCTCTCTGG 682
Db 3970 TTTATGTCGTCGTCGTAATGGAACACTGAAATCTGCAATATCCATAG 3929

RESULT 15
US-09-641-638-144

; Sequence 144, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GENSET.051CPI
; CURRENT APPLICATION NUMBER: US/09/641.638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502.330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133.200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275.267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119.917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 144
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 12-734-201 : polymorphic base T or C
; NAME/KEY: misc_binding
; LOCATION: 481..500
; OTHER INFORMATION: 12-734-201.misl, potential
; NAME/KEY: misc_binding
; LOCATION: 502..521
; OTHER INFORMATION: 12-734-201.misl2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 582..701
; OTHER INFORMATION: upstream amplification primer, complement
; NAME/KEY: primer_bind
; LOCATION: 301..320
; OTHER INFORMATION: downstream amplification primer
; NAME/KEY: misc_binding
; LOCATION: 489..513
; OTHER INFORMATION: 12-734-201 potential probe
US-09-641-638-144

Query Match      7.2%; Score 54.2; DB 4; Length 1001;
Best Local Similarity 66.5%; Pred. No. 1.5e-05;
Matches 115; Conservative 0; Mismatches 43; Indels 15; Gaps 2;

QY 533 TATTTTATGAGCAATGCTATTTATGCAATTAATAGACTCCAGTATTTGTAACATAAC 592
Db 7 TCTTTTATGACATAATGCTATTTATACACTTAAT-----ATAATATATCAGACT 56

QY 593 TTTAAATGCACTGGGAGATAAAGT-----ATTGCTCTTTTATGATATTTGCTTTATG 647
Db 57 TTTATATGCAATGGGAAACCAAAATGCTGACTTGCATTATGCTATATTTGCTTTATG 116

QY 648 CAGTAGTCTGTAATGGAACACTACATTATCTCTGCTGGTACACCTGTATACAGAA 700
Db 117 TGAATGCTGGAACCGAACAGTAGTATCTCCGAGGTATGCTGTATATGGA 169

Search completed: April 11, 2003, 22:47:55
Job time : 260.738 secs
```



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 21:30:39 ; Search time 48.0484 Seconds  
(without alignments)  
13710.183 Million cell updates/sec

Title: US-09-914-152-3\_COPY\_2000\_2750

Perfect score: 751  
Sequence: 1 cctccttggtgctcccaag.....tgatggaggctggaagtc 751

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 593429 seqs, 43858390 residues

Total number of hits satisfying chosen parameters: 1186858

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_applications\_NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	330.8	44.0	14117	9	US-09-954-531-1377
2	305.8	40.7	143306	10	US-09-729-920-3
3	301	40.1	460	10	US-09-954-456-1096
4	294.6	39.2	143068	10	US-09-967-768A-316
5	280.8	37.4	147309	10	US-09-742-312-3
6	273.2	36.4	489	10	US-09-864-761-16010
7	270	36.0	705	10	US-09-867-701-10850
8	258.2	34.4	402850	9	US-09-844-653-5
9	256.6	34.2	186957	9	US-10-185-770-3
10	251	33.4	1503841	9	US-09-946-807-1
11	251	33.4	1503841	10	US-09-795-668-1
12	251	33.4	1503841	10	US-09-795-668-1
13	249.8	33.3	476	10	US-09-864-761-10783
14	247.2	32.9	32176	9	US-10-091-504-1953
15	247.2	32.9	32176	10	US-09-764-869-1953
16	247.2	32.9	32178	9	US-10-091-504-1954
17	247.2	32.9	32178	10	US-09-764-869-1954
18	247.2	32.9	32191	9	US-10-091-504-1955
19	247.2	32.9	32191	10	US-09-764-869-1955

20	241.8	32.2	470	10	US-09-864-761-1095	Sequence 1095, Ap
21	237.2	31.6	28897	10	US-09-764-877-3897	Sequence 3897, Ap
22	224.8	29.9	514	10	US-09-864-761-8135	Sequence 8135, Ap
23	222.2	29.6	415	10	US-09-864-761-972	Sequence 972, App
24	221	29.4	459	10	US-09-864-761-609	Sequence 609, App
25	220.8	29.4	65359	10	US-09-804-472-3	Sequence 3, Appli
26	220	29.3	508	10	US-09-864-761-7194	Sequence 7194, Ap
27	212.6	28.3	592	10	US-09-864-761-16806	Sequence 16806, A
28	212	28.2	480	10	US-09-864-761-1910	Sequence 1910, Ap
29	211.2	28.1	484	10	US-09-864-761-467	Sequence 467, App
30	201.2	26.8	470	10	US-09-867-701-1599	Sequence 1599, Ap
31	199.6	26.6	538	10	US-09-864-761-8343	Sequence 8343, Ap
32	195.4	26.0	577	10	US-09-864-761-6667	Sequence 6667, Ap
33	195	26.0	472	10	US-09-864-761-5421	Sequence 5421, Ap
34	186.4	24.8	485	10	US-09-864-761-21750	Sequence 21750, A
35	177.2	23.6	461	10	US-09-864-761-6089	Sequence 6089, Ap
36	177	23.6	532	9	US-10-091-572-730	Sequence 730, App
37	177	23.6	532	9	US-10-091-572-732	Sequence 732, App
38	171.2	22.8	560	10	US-09-864-761-13192	Sequence 13192, A
39	170.6	22.7	419	10	US-09-864-761-3452	Sequence 3452, Ap
40	169.8	22.6	414	10	US-09-864-761-14183	Sequence 14183, A
41	168.2	22.4	472	10	US-09-864-761-473	Sequence 473, App
42	167.2	22.3	560	10	US-09-864-761-12852	Sequence 12852, A
43	165.2	22.0	473	10	US-09-864-761-4014	Sequence 4014, Ap
44	164.6	21.9	449	10	US-09-867-701-7289	Sequence 7289, Ap
45	164.2	21.9	472	10	US-09-864-761-4171	Sequence 4171, Ap

ALIGNMENTS

RESULT 1  
US-09-954-531-1377  
; Sequence 1377, Application US/09954531  
; Patent No. US20020165180A1  
; GENERAL INFORMATION:  
; APPLICANT: Weaver, Zoe  
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Usin  
; FILE OF INVENTION: Gene Sets  
; FILE REFERENCE: 689290-77  
; CURRENT APPLICATION NUMBER: US/09/954,531  
; CURRENT FILING DATE: 2002-05-02  
; PRIOR APPLICATION NUMBER: US/60/233,133  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: US/60/234,009  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: US/60/234,034  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: US/60/234,509  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: US/60/234,567  
; PRIOR FILING DATE: 2000-09-22  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1377  
; LENGTH: 14117  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-954-531-1377

Query Match 44.0%; Score 330.8; DB 9; Length 14117;  
Best Local Similarity 73.6%; Pred. No. 6e-72;  
Matches 493; Conservative 0; Mismatches 162; Indels 15; Gaps 5;

QY	36	GATGAGCCACCATGTCGACCTTCATGATAAACTTCAGTGATGAGAGCTGCCTCTTA- 94
Db	439	GATGTCATGATGTTGTCCTATGTTATGCCAAACGACGATGAGAACTGCTTTAT 498
QY	95	TGATGAACAAAGAGTGGTTCTTTGAAATGGAATCTCTCTGGTGAAGATGCTGTGAA 154
Db	499	TCATGAACAAAGAAAGTGGTTCTTTGAGATGGAACATCTCTCTGGTGAAGATGCTGTGAA 558
QY	155	CATTGTTGAATGACAGAAGAATTTACAGTTGTACATAGAGTTAGTTGATGAAGCAGT 214

Db 559 CATCACTGAAATGGCAACAAGGATTTGGAAATACACTACATCAATTTAGTTAATAAGGCAGT 618  
QY 215 AGCAGGATTCGAGAGGATCGCATTTCCAAATTTCAAATAAGTTCTTCTGTGGTAAAGTCT 274  
Db 619 GCGAGGTTTGAGAGGACTGACCCCTAGTTTGAAGAAATTTCTACTGT---TAAATGCT 675  
QY 275 ATCAAAATGGCGTGGCTGCTACAGAGAAATCTATCATGAAAGGAAGAGTCAATTTGATGTG 334  
Db 676 ATCAAAACAGCATCATGCTACAGGGAATCTTTTCATGAAAGGAAGAGTCAACTGATGCG 735  
QY 335 CCAAACTTCATGTTGCTGCTATTTTAAGAAATTTGTCAGGACCCCACTTCAACAAC 394  
Db 736 CCAAACTTCATGTTGCTGCTATTTTCAGAAATTTGCCACAGCCACACTAACTTGCAGCAAT 795  
QY 395 CATGACCTGATCAGTCAGGAGCCATCCACATTTAGGCGAGAACCTCCAGCAGTAAAG 454  
Db 796 CATGACCTGATCAGTCCTCAGCCATCAATTTAGGCGAGAACCTCCAGCAGTAAAG 855  
QY 455 ATTATGATTCCTAAAGGATCAGTGAACATTTAGCATTTTAAAGCAATTAAGTATTTT 514  
Db 856 ATTATGATTCCTAAAGGATCAGTGAACATTTAGCATTTTAAAGCAATTAAGTATTTT 914  
QY 515 TAGCTAAGATATGATGTTATTTTATAGGCATTAATGCTATTTATGATTTTAAAGTATTTT 574  
Db 915 AAATTAAGTTATATACATTTGTTTATAGGCATTAATGCTATTTATGATTTTAAAGTATTTT 974  
QY 575 GTATATTTGATAACATAA-CITTTAAATGCACTGGGAGATAAAGTATTTG-----CT 624  
Db 975 GTATAGTGAATATATGTTTACATACACAGAAACCAAAAGTTTGTGTGACTCACT 1034  
QY 625 CTTTATGATATTTGCTTTATTTGAGTGTCTGTATGAAACTACATTTATCTCTTGGT 684  
Db 1035 TTATGGTTACATTTGCTTTATTTGTTGTTGCTGTAGAACTGAACATGCACTATCTGTAGGT 1094  
QY 685 ACACCTGTAT 694  
Db 1095 ATGCTCTGAT 1104

RESULT 2  
US-09-729-920-3  
; Sequence 3, Application US/09729920  
; Patent No. US20020103115A1  
; GENERAL INFORMATION:  
; APPLICANT: GUEGLER, Karl et al  
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,  
; TITLE OF INVENTION: AND USES THEREOF  
; FILE REFERENCE: CL000858  
; CURRENT FILING DATE: US/09/729,920  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 143306  
; TYPE: DNA  
; ORGANISM: Human  
US-09-729-920-3

Query Match 40.7%; Score 305.8; DB 10; Length 143306;  
Best Local Similarity 74.0%; Pred. No. 3.3e-65;  
Matches 485; Conservative 0; Mismatches 152; Indels 18; Gaps 7;

QY 47 ATGGTCGACTTCATGATAAAACCTTCAGTGGATGAGGAGCTGCTCTTAT-GATGAACAAA 105  
Db 88548 ATGCTGCATCTCATGATAAAACCTTCAGTGGATGAGGAGCTTCTTCTTATGATGACAAA 88607  
QY 106 GAAGGTGTTTCTTGAATGAATCTACT----CCTGTGAAGATGCTGTGAACATTTGT 161  
Db 88608 GAAAGTGATTTCTTGAATGAATCTACTCTCTGCTGTGAAGATGCTGTGAACATTTGT 88667  
QY 162 GAAATGACAAGAAATTTACATGTTTACATAGATTTAGTTGATGAACCACTAGCAGGA 221

Db 88668 GAAATGACAACAAGAAATTCAGAATATTTATATCAACTTAGTTGATTAAGCACTGCGAGG 88727  
QY 222 TTCGAGAGGATCGATTTCCAAATTTCAAATAAGTTCTTCTGTGGTAAAGTATGCTATCAAAAT 281  
Db 88728 TTTGAGAAGATTTGATTTCCAAATTTGAAAAA- AAGTTTCTACTGTAGGTAAATGCTATCTAAC 88786  
QY 282 GCGTGGCATGCTACAGAGAAATCTATCATGAAAGGAAGAGTCAATTTGATGTGGCAACT 341  
Db 88787 AGCATTTGCAATTTACAGAGAA-----ATTGAAAGGAAGAGTCAATTTGATGTGGCAACT 88842  
QY 342 TCATTTGTTGCTGATTTTAAAGAAATTTGTCAGGAGCACCCCAACCTTCAACAACCATGACC 401  
Db 88843 TCTTTGTTGCTGATTTTAAAGAAATTTGTCAGGAGCACCCCAACCTTCAACAACCATGACC 88902  
QY 402 CTGATCAGTCAGGAGCATCCCAATTTGAGGCGAGAACCTCCAGCAGTAAAGATTTATGA 461  
Db 88903 CTGATCAGTCAGTGTAGTATCATCATCAATCAAGAGAACCTCCGCGCAAGCAATTTATGA 88962  
QY 462 TTCTCTAAAGGATCAGATGAACATTTAGCAATTTTAAAGCAATTAAGTATTTTACGTAA 521  
Db 88963 CTGTTGAAACACTCAGATTTATTTAGAAATTTTGGTAATAAGTATTTTCAAGTTAAGG 89022  
QY 522 GATATGATTTGATTTTATTTAGGCATTAATGCTATTTATGCAATTTAATAGACTCCAGTATTT 581  
Db 89023 AGTATATAT-ACATTTTATGACCTATGCTATTTGTACACTCAATAGACTGTAGTAGACT 89081  
QY 582 GTAACATTAAC-TTTAAATGCACTGGGAGATAAAGTATTTTGTCTCTTTTATGATATTTGC 640  
Db 89082 GTAAACATAACTTTTATATGCACTGGGAAACAAAA-----AAGAAATGTGTGACTCAT 89135  
QY 641 TTTATTTGCACTAGTCTGTAATGGAACACTACATTTCTCTTGGGTACACCTGTATA 695  
Db 89136 TTTATTTGCTGTCTGTAAGTCTAACTCACAATGTCTCTGACGTATGCTGTATA 89190

RESULT 3  
US-09-954-456-1096/c  
; Sequence 1096, Application US/09954456  
; Patent No. US20020115057A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Paul  
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using  
; TITLE OF INVENTION: Sets  
; FILE REFERENCE: 689290-76  
; CURRENT APPLICATION NUMBER: US/09/954,456  
; CURRENT FILING DATE: 2001-09-18  
; PRIOR APPLICATION NUMBER: US/60/233,617  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: US/60/234,052  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: US/60/234,923  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,134  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,637  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,638  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,711  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,720  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,840  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,863  
; PRIOR FILING DATE: 2000-09-27  
; NUMBER OF SEQ ID NOS: 2276  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1096  
; LENGTH: 460  
; TYPE: DNA  
; ORGANISM: Homo sapiens



Db 103535 TAGTTCTTCAGAGT-----CTACTGGTGAAGTGGCGTGAACATGTTTAAATGATG 103483  
QY 171 AGAAGAATTTACAGTGTACATAGAGTTAGTTGATGAAGCAGTAGCAGGATTCGAGAGG 230  
Db 103482 ACAAGGATTTAAATATCTGTAAACTTAGTTGACAAAGCAGCAGCAAGGCTTGAGGTT 103423  
QY 231 ATCCATTCCAAATTTCAAAATAGTTCTTCTGTGGTAAATGCTATCAATGCGTCGCA 290  
Db 103422 ATTGAGTTTAATTTTGAAGAAGTTCTACTTTGGGTCAATGCTATGAAACTGCATTGCA 103363  
QY 291 TGCTACAGAGAAATCTATCATGAAGAGAGAGTCAATTCATGTGGCAAACTTCATTGTTG 350  
Db 103362 TGCTACAGAGAGATCTTTAGTGAAGAGAAAGTTAATTCAGTGAATTAATTTATTG--- 103306  
QY 351 TCGTATTTTAAAGAAATTTCTCAGGACACCCCAACCTTCAACCACTGACCTGATCAGT 410  
Db 103305 -----TCTTAAGAAATTTGCCAGGCGCACTTTAACTTCAACCACTGATCAGT 103251  
QY 411 CAGGAGCAATCCACATTCAGGCGGAGAACTCCAGCAGTAAAGATTTGATTTCTCTAAA 470  
Db 103250 CAGTGGCCATCAACATTTAGGCAAGACCCCTCCACCAGCAAAAGATTATGCCTTGTGAA 103191  
QY 471 GGATCAGATGAACATTTAGCAATTTTAAAGCAATAAGTATTTTACGTAAAGATGAT 530  
Db 103190 GGCTCAGATGATTTAGCAATTTCT--AGCAATGAGTAATTTA-AAATATGTTCT 103134  
QY 531 GTATTTTAAAGCAATGATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 590  
Db 103133 CACTTTTTTAAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 103074  
QY 591 AC-TTTAAATGCACTGGGAGATAAA-----GTATTTGCTCTTTTATGATATTTGC 640  
Db 103073 ACTTTTCTGCACTGGGAAACCAAAATTTCTCGGACTTTCTTTGTTGGGATCTTGT 103014  
QY 641 TTATTTTCAGTGTCTGTAATGGAACATACATATCTCTTGGGTACACCTGTATACAGAA 700  
Db 103013 TTGTTTGTGGTCTGAAACTGAACTGCAATGCAATTCCTTGAGATATGCTGTATTAAT 102954  
QY 701 AGAAATTT 708  
Db 102953 GTAACATAT 102946

RESULT 6  
US-09-864-761-16010/c  
; Sequence 16010, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aemolica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; PRIOR FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 16010  
; LENGTH: 489  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC010888.2  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.99  
US-09-864-761-16010

Query Match 36.4%; Score 273.2; DB 10; Length 489;  
Best Local Similarity 78.7%; Pred. No. 2.1e-58;  
Matches 389; Conservative 0; Mismatches 98; Indels 7; Gaps 5;

QY 69 TTCAGTGGATGAGGAGTGGCTCTTAT-GATGAACAAGAGAGTGGTTCTTGAATGGA 127  
Db 489 TTCATGATGGAGTGGAGTGGCTTTTATGATGAGCAAGAAATGGTTTATGAGTGA 430  
QY 128 ATCTACTCCTGTGGAAGATGCTGTGAACATTTGTGAATGACAAAGAAATTTACAGTG 187  
Db 429 ATCTACTTCTGTGGAAGATGCTATGAACATTTGTGAATGACAAATACAGATATAACTT 370  
QY 188 TTACATAGAGTTAGTTGATGAAGCAGTAGTACAGAGATTCGAGAGATCGATTCCAAATTTCAA 247  
Db 369 GTACATAAACTTAGTTAAGAAAGCAATGCGAGAGTTTGAGAGGATTCACATTTCAATTTCAA 310  
QY 248 AATAAGTTCTTCTGTGGTAAATGCTATCAATGCGTGGCAGTGTACAGAGAAATCTA 307  
Db 309 AAGAAGTTCTTCTGCTCACTCAATGCTATCAAAACAGCATACATGCTACGAGATACCTT 250  
QY 308 TCATGAAGAGAGAGTCAATTTGATGTGGCAAACTTCATTTGCTGTTATTTAAGAAAT 367  
Db 249 TTGTGAAGAGAGAGTCCATTTGATGAGCAAACTTCATTTGCTGTTATTTATGAATTT 190  
QY 368 GTCAGGACCAACCCCAACCTTCAACCAACCATGACCTGATCAGTCAGGAGCCATCCACATT 427  
Db 189 GCCACACCAACCCCAACCTTTCAGCAACCAACCCGCTGCTGATCACTCAACCAACCAT-CACT 131  
QY 428 GAGCGAGAGACCTCCAGCAGTAAAGATTTATGATTTCTCTAAAGGATCAGATGAACATTA 487  
Db 130 GAAGAAAGATTTCTTACCAGGAAAGAGATGATGAT-TCTGAAGACTCAGATATCGTTA 73  
QY 488 GCATTTTTTTAAGCAATAAGATATTTT-ACGTAGATATGATGTTATTTTATTTAGGCAT 546  
Db 72 GTATGTTT--AGCAATAAGAAATTTTTTAAATTAAGTATCTATGTTGTTTTTTAGGCAT 15  
QY 547 AATGCTATTATGCA 560  
Db 14 AATGTTATTTTACA 1

## RESULT 7

```

US-09-867-701-10850/c
; Sequence 10850, Application US/09867701
; Patent No. US2002013237A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10850
; LENGTH: 705
; TYPE: DNA
; ORGANISM: Homo sapien
;
; FEATURE:
;
; NAME/KEY: misc_feature
; LOCATION: (1)..(705)
;
; OTHER INFORMATION: n = A,T,C or G
US-09-867-701-10850

```

Query Match	36.0%	Score 270;	DB 10;	Length 705;
Best Local Similarity	70.6%	Pred. No. 1.6e-57;		
Matches 430; Conservative	0;	Mismatches 165;	Indels 1	

5

Qy	96	GATGAACAAGAGGTGGTTTC	TTGAAATGGAATCTACTCTG	TGGAAGATGCTGTGAAAC	155
Db	687	GATGAGCAAGAAGAAATGAT	TTTTTGAATGAATATACTCT	CGGTGGAGATGCCATGACC	628
Qy	156	ATTGTGTGAATGACAAGAAG	ATTTACAGTGTTCACATAG	AGTTAGTTGATCAACGCAGTA	215
Db	627	ATTGTGCAATGCAAGAAG	ATTTAGATAGTACATA-ANC	TCAATTTGATACACGCAGTA	569
Qy	216	GCAGGATTCGAGAGGATCG	ATTCCTCAATTAAGTTTCT	CTGTGGGTAATAATGCTA	275
Db	568	GCAGGGTTTTAAGAGGAT	TGACTCCAATTNTAAAGA	AGAGTTCTACTGTGAGTANAATGCTA	509
Qy	276	TCAAATGGCGTCGCTAC	AGAGAATCTATCATGAA	AGGAAGAGTCAATTTGATGTGG	335
Db	508	TCACACAGCATTTGTG	CTACAGAAATCTTTTGG	AAAGGAAGAGTCAATCAATACAG	449
Qy	336	CAAACTTCAATTTGTG	CTGATTTTAAAGAAAT	TGTCAGAGCACCCCAACCTTCAACAACC	395
Db	448	CAAACTTTATTTGTT	ATTTTATTTTAAAGAG-TC	CCAAAGCCACCCACCTTTTCAGCAGCC	390
Qy	396	ATGACCCCTGATCAGT	CAGAGCCATCCATTTG	AGGGGAGAACTCCAGCAGTAAAAAGA	455
Db	389	ACCAACTTGTCAATC	AGCAGGTATCCGCACT	GAGGTAAGACTCTTCATCAGCAAAAAGA	330
Qy	456	TTATCATTTCTTAAAG	GATCAGATGAACATTT	TTTATGACCAATAAGTATTTTT	515
Db	329	CTGTAACTCATTC	AGGCTCTAATAATCAT	TATGATCTTTTT--AGCAATAAAATATTTTA	272
Qy	516	ACGTAAAGATATGT	TATTTTATTTAGGCATA	TATGCTATTTATGATTAATAGACTCCAG	575
Db	271	ATTCAGATTTTGTAT	ATTTTATTTTAGACATA	TATGCTATTTGCACATTTAATGACAAATAG	212
Qy	576	TATATTGTAACATAA	-CTTTAAATGCATCGG	AGATATAAAAGTAT-----TTGCTC	625
Db	211	TATAGTATAAGCAT	ATAATTTTATATGC	ACTGGAATAATTTATGTAACCTTGTTT	152
Qy	626	TTTTATGATATTTG	CTTTATTCGACTAGT	CTGTPAATGGAACACTACATATCTCTTGGGTA	685
Db	151	TATTGCATATTTGG	CTTTGTGCACAGTGG	TCGTGGAATCAAACTCGCAATATCTTAGAAGTA	92
Qy	686	CACCTGTAT	694		
Db	91	TGCCTGTTT	83		

RESULT 8

```

US-09-844-653-5
; Sequence 5, Application US/09844653
; Publication No. US2003005437A1
; GENERAL INFORMATION:
; APPLICANT: Richards, Julia
; APPLICANT: Rozsa, Frank
; TITLE OF INVENTION: Detecting and Treating Eye Disease
; FILE REFERENCE: UM-06105
; CURRENT APPLICATION NUMBER: US/09/844,653
; CURRENT FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 402850
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (46565)..(46565)
; OTHER INFORMATION: The n at this position can be a, c
; NAME/KEY: misc_feature
; LOCATION: (117443)..(118256)
; OTHER INFORMATION: The n at these positions can be a,
; NAME/KEY: misc_feature
; LOCATION: (118272)..(118736)
; OTHER INFORMATION: The n at these positions can be a,
; NAME/KEY: misc_feature
; LOCATION: (120172)..(120640)
; OTHER INFORMATION: The n at these positions can be a,
; NAME/KEY: misc_feature
; LOCATION: (122654)..(122654)
; OTHER INFORMATION: The n at this position can be a, c
; NAME/KEY: misc_feature
; LOCATION: (203527)..(203963)
; OTHER INFORMATION: The n at these positions can be a,
; NAME/KEY: misc_feature
; LOCATION: (215340)..(215340)
; OTHER INFORMATION: The n at this position can be a, c
; NAME/KEY: misc_feature
; LOCATION: (220202)..(220650)
; OTHER INFORMATION: The n at these positions can be a,
; NAME/KEY: misc_feature
; LOCATION: (222415)..(222416)
; OTHER INFORMATION: The n at these positions can be a,
; US-09-844-653-5

```

Query Match	34.4%	Score 258.2;	DB 9;	Length 402850;
Best Local Similarity	70.2%	Pred. NO. 3.8e-53;		
Matches 453; Conservative		0; Mismatches 168;	Indels 24;	Gaps 7;

QY	52	CGAC	TT	CAT	GAT	AAAA	CT	TCAG	TGG	ATG	AGG	AGC	TG	CT	CTTA	-	TG	AT	GA	ACA	AGA	AGG	110	
Db	342633	CAAT	CT	CAT	CA	TCA	AT	TTA	TAT	GTA	AG	AT	TG	CT	TCT	-	TG	AT	GA	ACA	AGA	AGG	342692	
QY	111	TGG	TT	CT	TT	GA	AT	GAA	T	GAA	T	CT	CT	G	T	G	A	G	A	T	G	C	T	170
Db	342693	TGG	TT	CT	TT	GA	AT	GAA	T	CT	CT	G	T	G	A	T	G	C	T	G	C	A	T	342752
QY	171	AGA	AA	GA	AT	T	T	C	A	G	T	T	T	C	A	T	T	G	A	T	G	A	G	230
Db	342753	ACA	A	AG	GA	AT	T	T	A	G	A	T	T	T	A	T	-	A	T	T	G	T	G	342810
QY	231	AT	C	G	A	T	T	C	CA	AA	T	A	A	G	T	T	C	T	C	T	G	T	G	290
Db	342811	AT	T	G	A	C	T	T	C	A	G	T	T	C	A	A	T	A	A	G	T	T	C	342870
QY	291	TG	C	T	A	C	A	G	A	A	T	C	A	T	C	A	T	T	G	A	T	G	T	350
Db	342871	TAC	T	A	C	A	G	A	A	T	C	T	T	C	A	G	A	A	G	A	G	A	G	342930
QY	351	TC	G	T	A	T	T	T	A	G	A	A	T	T	C	A	G	A	G	A	C	C	C	409

```

Db 342931 TCTTATTTTAAAGAAATTTCCATAGTCCCATTAACCCCTTAGCAACTCATACCCCTGATCAT 342990
QY 410 TCAGAGCATCCACATTTAGGCGGAGAACCTCCAGCAGTAAAGAAATTTATGATTCCTCTAA 469
Db 342991 TCAGCAGCTTTTAACTACCCAGACAGTCTCTCCATCA -CAAAAATATTAGAACTCCTGTA 343049
QY 470 AGGATCAGATCAACATTAGCATTTTAAAGCAATAAAGTATTTTACGTAGATATGTA 529
Db 343050 AGGCTCAGAATATGGTAAC--TTTTTTTAAACAATAAAGTATTTTAAATCAA----- 343100
QY 530 TGTATTTTTTAGCATTAATGCTATATGATTTTAAATAGACCCAGTATATTGTAACAT 589
Db 343101 -GGTATGATAGTGTTTTACTCTCTATGGTGGCTTTAATAGACTACATAATATGTAACAT 343159
QY 590 AACTTTAAATGCACCTGGGAGATAAAGTATTTGCTTTTATGATATTTGCTTTATTGCA 649
Db 343160 AACTTTTA-----TACCACCTGGGAATCCCAATTTGTGACTCATTATTGTCG 343210
QY 650 GTAGTCTGTAAATGGAACACTACATTTATCTCTGGGTACACCTGTAT 694
Db 343211 ATGGTCTGGAACCTGAGCTGCAATCTCTCAGGTATGCTGAT 343255

RESULT 9
US-10-185-770-3/c
; Sequence 3, Application US/10185770
; Publication No. US2003002217A1
; GENERAL INFORMATION:
; APPLICANT: CECARDI, Toni et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL0001247
; CURRENT APPLICATION NUMBER: US/10/185.770
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 50/301.852
; PRIOR FILING DATE: 2001-07-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 186957
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(186957)
; OTHER INFORMATION: n = A,T,C or G
US-10-185-770-3

Query Match 34.2%; Score 256.6; DB 9; Length 186957;
Best Local Similarity 69.6%; Pred. No. 6.4e-53;
Matches 460; Conservative 0; Mismatches 169; Indels 32; Gaps 7;

QY 52 CGACTTCATGATAAACTTCAGTGGATGAGAGCTGCTCTTATGATGAACAAAGAGGT 111
Db 46694 CAATTTCATGATAAACTTGAATGGATGAGAGTGTCTCTTATGGATAGCAAGAAAG 46635

QY 112 GCTTTCTGAAATGAATCTACTCTCTGGTGAAGATGCTGTGAACATTTGTAATGACAA 171
Db 46634 TGTTCCTTCAGATGAATCTACTCTCTGGTGAAGATGCTATAACATCACCAGAACACAA 46575

QY 172 GAAGAATTTTACAGTGTACATAGATTAGTATGCA-----AGCAGTACAGATTGCA 226
Db 46574 CAGATGATTTGATATTTACATAAATGATGATTAATTAAGCAGTGGCAAGGTTGA 46515

QY 227 GAGATCGATTCCTCAATTTCAAAATAGTTCTTCTGTGGGTAAATGCTATCAATGCGGT 286
Db 46514 CAGGATTCGACTCCGATTTCTAAAGAAATGCTACTATGATGATAGACATATCAACACAGAA 46455

QY 287 CGC-ATGCTACAGAAATCTATCATGAAGGAAGAGTCAATTTGATGTGGCAACTTCAT 345
Db 46454 CACAATGCTACAGATAAACAATTTCATTAAAGGAAGAGTCCATCAATGTGGCAAAACATTGT 46395
```

```

QY 346 TGTTCGTGATTTTAAAGAAATTTGTCAGGACCCACCCCAACCTTCAACCAACCATGACCCCTGA 405
Db 46394 TGTCTTC--TTTTAAAAAAGCTGCCACGACCCCAACCTTCCAGCAACCAACCACTGA 46337
QY 406 TCAGTCAGAGCCATCCACATTTAGGCGGAGAACCTCCAGCAGTAAAGAAATTTATGATTCCT 465
Db 46336 TCAGTCAGCTGCCATCAACATCAAGCAAGACCCCTCTACCAGCAACAGATATGATTCATCA 46277
QY 466 CTAAGGATCAGATGAACATTAG-----CATTTTTTAAAGCAATAAAGTA-TTTTTACGT 519
Db 46276 GTGAAGGCTCAGATGATTTAGTGGGTTTTTTTTTTAGCAATAAAGTACTTTTTTAATT 46217
QY 520 AAGATATGCTATGTTATTTT-TTTAGGCATATGCTATTTATGCACTTTTAATAGACTCCAGTAT 578
Db 46216 AAGAAATGCTATGTTTATTTTCGACATAAAGCTATTTGGCTACTTAAATAGACTACAGCAC 46157
QY 579 ATTGTAACATAAATTTTAAATGCACCTGGGAGATAAAGTATTTGCTCTTTTATGATTTT 638
Db 46156 AGTGTAAACATA-----TGCAAGAAAGAACCAAAAATTTGTGTGACTC 46114
QY 639 GCTTTATTGCACTGCTGTATGTAATGGAACACTACATATCTCTGGGTACACCTGTATACAG 698
Db 46113 ACTTTACTGTGTGCTGCTGGAACACACACCCACATATCTCTGAAGTATGCCTGTATTTAT 46054
QY 699 A 699
Db 46053 A 46053

RESULT 10
US-09-946-807-1
; Sequence 1, Application US/09946807
; Patent No. US2002016514A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn
; APPLICANT: Steinthorsdottir, Valgerdur
; APPLICANT: Gulcher, Jeffrey R.
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345.2004-001
; CURRENT APPLICATION NUMBER: US/09/946.807
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US/09/795.668
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/515,716
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1503841
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: r=g or a
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: y=t/u or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: k=g or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: s=g or c
; FEATURE:
```

NAME/KEY: misc\_feature  
LOCATION: (1)...(1531)  
OTHER INFORMATION: w=a or t/u  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(1531)  
OTHER INFORMATION: d=g or c or t/u  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(1531)  
OTHER INFORMATION: d=a or g or t/u  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(1531)  
OTHER INFORMATION: h=a or c or t/u  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(1531)  
OTHER INFORMATION: v=a or g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(1531)  
OTHER INFORMATION: n=a or g or c or t/u

US-09-946-807-1  
Query Match 33.4%; Score 251; DB 9; Length 1503841;  
Best Local Similarity 69.2%; Pred. No. 4.7e-51;  
Matches 448; Conservative 0; Mismatches 180; Indels 19; Gaps 7;  
QY 52 CGACTTCATGATAAACCTTCAGTGGATGAGGAGTGCCTCTTATGATGAACAAAGAGGT 111  
Db 487762 CAATCTCATGCTAAACCTGAATGATTGAGCAATGTTTCTTACAGATGAGCAA-AATAA 487820  
QY 112 GGTTCCTTGAATGGATCTACTCTCGTGAAGATGCTGTGAACATTTGTTGAATGACAA 171  
Db 487821 GTTTCTTGAGTGGATCTGCTCTGATGAATCTCTGTGAACATTTATTTGAATGACAA 487880  
QY 172 GAAAGAAATTTACAGTGTACATAGAGTTAGTTGATGAAGCAGTACGAGATTCGAGAGA 231  
Db 487881 CAAAGGATTAGACATTACATAAAGTGGTGTAGATAAGCAGACAGGGTTTGAAGGA 487940  
QY 232 TCGATTCCAAATTTCAAATTAAGTTCTTCTGGGTAAATGCTATCAAAATGGCGTCGAT 291  
Db 487941 TTGCTTCCAAATTTCTCAAGAAATTTTATTTGGGTAGCATGCTATCAAGAGCATATAT 488000  
QY 292 GCTACAGAGAAATCTATCATGAAGGAGAGTCAATTTGATGTGGCAAACTTCATTTGT 351  
Db 488001 GCCACAGAGAAATATTTTATAAAGGAAGAG- - -TTAAAGCAGCAAAATTTTCATTGT 488056  
QY 352 CGTATTTTAAAGAAATTTGTCAGGACCCACCCCAACCTTCAACACCATGACCTGATCAGTC 411  
Db 488057 CTTTAAAGAAACNATTAGCACGCCATCTCAGACTTCAGCAACCCATCCTGATCAGTC 488116  
QY 412 AGGAGCATCCACATTCAGGCGAGACCTCCAGCAGTAAAGATTTATGATCTCTTAAG 471  
Db 488117 AGCAGCATCAACATCAAGGAGAGCCCTCCACAGCAAAAGATTTATGACTTTCTGAAG 488176  
QY 472 GATCAGATGAACATTAGCATTTTAAAGCAATAAAGTATTTTACGTAAGATGTATG 531  
Db 488177 ATTTCAGATGATCATTTAGCATTTT- - -AGCAATATGATTTTACA- - -AATTAAGGTATG 488230  
QY 532 TTATTTTTCAGTAAATGCTATTTATGCAATTAATAGACTCCAGTATATTGTAACATAAA 591  
Db 488231 TACATTTTTCAG- - -ATATGTTTATGACACTTAATAGACTACAGTATAGTACAACTTTG 488288  
QY 592 CTTTAAATG- - -CACTGGGAGATAAAGTATTTTCTCTTTTATGATATTGCTTTAT- - 645  
Db 488289 TATGCACTGGGAAACCAAAAATTTAGTGACCTGTTTATTTATGATGTTTCACTTTATTA 488348  
QY 646 TGCAGTGTCTGTAATGGAACATACATTTATCTCTGGGTACACCTGT 692  
Db 488349 TTCAGTTTCACTAGAACTGAACCTGTATATCCCTAAGGTATGCTTAT 488395

RESULT 11  
US-09-795-668-1  
Sequence 1, Application US/09795668  
Patent No. US20020045577A1  
GENERAL INFORMATION:  
APPLICANT: Stefansson, Hreinn  
APPLICANT: Steinhorsdottir, Valgerdur  
APPLICANT: Guicher, Jeffrey R.  
TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE  
FILE REFERENCE: 2345.2004-001  
CURRENT APPLICATION NUMBER: US/09/795,668  
CURRENT FILING DATE: 2001-02-28  
PRIOR APPLICATION NUMBER: US 09/515,716  
PRIOR FILING DATE: 2000-02-28  
NUMBER OF SEQ ID NOS: 1531  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 1503841  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(1531)  
OTHER INFORMATION: r=g or a  
NAME/KEY: misc\_feature  
LOCATION: (1)...(1531)  
OTHER INFORMATION: y=t/u or c  
NAME/KEY: misc\_feature  
LOCATION: (1)...(1531)  
OTHER INFORMATION: m=a or c  
NAME/KEY: misc\_feature  
LOCATION: (1)...(1531)  
OTHER INFORMATION: k=q or t/u  
NAME/KEY: misc\_feature  
LOCATION: (1)...(1531)  
OTHER INFORMATION: s=g or c  
NAME/KEY: misc\_feature  
LOCATION: (1)...(1531)  
OTHER INFORMATION: w=a or t/u  
NAME/KEY: misc\_feature  
LOCATION: (1)...(1531)  
OTHER INFORMATION: b=g or c or t/u  
NAME/KEY: misc\_feature  
LOCATION: (1)...(1531)  
OTHER INFORMATION: d=a or g or t/u  
NAME/KEY: misc\_feature  
LOCATION: (1)...(1531)  
OTHER INFORMATION: h=a or c or t/u  
NAME/KEY: misc\_feature  
LOCATION: (1)...(1531)  
OTHER INFORMATION: v=a or g or c  
NAME/KEY: misc\_feature  
LOCATION: (1)...(1531)  
OTHER INFORMATION: n=a or g or c or t/u

US-09-795-668-1  
Query Match 33.4%; Score 251; DB 10; Length 1503841;  
Best Local Similarity 69.2%; Pred. No. 4.7e-51;  
Matches 448; Conservative 0; Mismatches 180; Indels 19; Gaps 7;  
QY 52 CGACTTCATGATAAACCTTCAGTGGATGAGGAGTGCCTCTTATGATGAACAAAGAGGT 111  
Db 487762 CAATCTCATGCTAAACCTGAATGATTGAGCAATGTTTCTTACAGATGAGCAA-AATAA 487820  
QY 112 GGTTCCTTGAATGGATCTACTCTCGTGAAGATGCTGTGAACATTTGTTGAATGACAA 171  
Db 487821 GTTTCTTGAGTGGATCTGCTCTGATGAATCTCTGTGAACATTTATTTGAATGACAA 487880  
QY 172 GAAAGAAATTTACAGTGTACATAGAGTTAGTTGATGAAGCAGTACGAGATTCGAGAGA 231  
Db 487881 CAAAGGATTAGACATTACATAAAGTGGTGTAGATAAGCAGACAGGGTTTGAAGGA 487940





Query Match	32.9%;	Score 247.2;	DB 9;	Length 32176;
Best Local Similarity	71.4%;	Pred. No. 5.5e-51;		
Matches 460;	Conservative	0;	Mismatches 153;	Indels 31; Gaps 9;
QY 74	TGGATGAGGAGTGCCTCTTAT	-GATGACAACAAGAGGCTGTTCTT	TGAATGCAATCTA	132
Db 14549	TGGATTAGGAGTGTCTTCTAT	TATGATGACCAATAAAGTGGTTC	TGTGAGAGAAATCTA	14490
QY 133	CTCCTGGTCAAGATGCTGTGA	CATTTGTAATGACAAGAAAGATTT	TACAGTGTTCACA	192
Db 14489	CTGCTGAGATGTTGTG-----	AATGTTGAAATTGACAC	-AAAGAATTCAGATATTACA	14437
QY 193	TAGATTAGTTGATGAAGCAG	TAGCAGGATTCGAGAGGATCGAT	TCCAAATTCAAAATAA	252
Db 14436	TAAACTTAGTTGNTAAGCAG	CAGCAAGTCTGCACAGGACTG	ACTCCAATTTTGA	14380
QY 253	GTTCTTCTGTGGTAAATGCT	TATCAAAATGGCGTGCATGCT	CACAGAAATCTATCATG	312
Db 14379	GTTCTGTGTGGCTTAAATG	CTATCAACAGCGCCGATGCT	TAAAGAGGCACTCTTTGTG	14320
QY 313	AAAGGAAGATCAATTGATGT	GGCAAACTTCATTGTTGCTG	TATTTTAAGAAATTTGCAG	372
Db 14319	AAAGGAAGATCAATCTGTG	GGGCAAACTTCATTGCTGTC	-----TTAAGAAATTTGCTGT	14265
QY 373	GACCAACCCCAACTTCAACA	CCATGCACTGATCAGT	CAGGAGCCATCCACATTTGAGGC	432
Db 14264	AGTCACCCCAACTTCAGCA	CCACTACCTGATTA	GTAGTCAGAGCCATCACAAGCAAAAC	14205
QY 433	GAGAACCTCCAGCAG	-TAAAAAGATTATGATTC	TCTTAAAGGATCAGATGAACATTTAGCAT	491
Db 14204	AAGACTATCCACCACAAAAA	AGATTATGATTCAC	TGATCACTGAAGGCTCAGATGC	14149

QY 492 TTTTAAAGCAATAAGTATTTTACGTAAGATATGTATGTTATTTTATGAGCATATGC 551  
 Db 14148 TTTTAAAGCAATAAGTATTTTATTTTAAAGTATGTACACGTTTATGATAGAACT 14089  
 QY 552 TATTATGCATTTAATAGACTCAGTATATTTAAACATAAC-TTTAAATGCACGTGGAGA 610  
 Db 14088 TATTGTACACATAGACTACAGTATTTGTGTAATAAACTTTTATATGCAGTGGAAA 14029  
 QY 611 TAAAA-----GTATTTGCTCTTTATGATATTTGTTTATGCACTAGTCTGTAAT 661  
 Db 14028 CCAAAAACTCATGAGACTAATTTTATGTGATAGTCTTTTATGTCAGCAGTTTGAAC 13969  
 QY 662 GGAACACTACATTTCTCTTTGGGTACACCTGTATACAGAAAGAAA 705  
 Db 13968 TGAACCTGCAATATCTCGTGGTATGCTGTACCCCTGAAAAA 13925

RESULT 15

US-09-764-869-1953/c  
 ; Sequence 1953, Application US/09764869  
 ; Patent No. US20020061521A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: PC007  
 ; CURRENT APPLICATION NUMBER: US/09/764,869  
 ; CURRENT FILING DATE: 2001-01-17  
 ; Prior application data removed - refer to PALM or file wrapper  
 ; NUMBER OF SEQ ID NOS: 2442  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1953  
 ; LENGTH: 32176  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-764-869-1953

Query Match 32.9%; Score 247.2; DB 10; Length 32176;  
 Best Local Similarity 71.4%; Pred. No. 5.5e-51;  
 Matches 460; Conservative 0; Mismatches 153; Indels 31; Gaps 9;  
 QY 74 TGGATGAGGAGCTGCCTCTTAT-GATGAACAAAGAGGTGTTCTTCAATGGAATCTA 132  
 Db 14549 TGGATTAGGAGTTGCTCTTATGGATGAGCAATTAAGTGGTTCTTGAAGAAATCTA 14490  
 QY 133 CTCCTGGTGAAGATGCTGTAACATTTGAAATGACAAGAAAGAAATTTACAGTGTACA 192  
 Db 14489 CTGCTGAAGATGTTGTG-----AATGTTGAAATGACAC-AAAGAAATTCAGAAATTTACA 14437  
 QY 193 TAGAGTTAGTTGATGAGCAGTAGCAGGATTCGAGAGATCGATTCCAAATTTCAAAATAA 252  
 Db 14436 TAACTTAGTTGATAAGCAGCAGCAAGTCTGACAGGACTGACTCCAATTTGA--AA 14380  
 QY 253 GTTCTTCTGTGGTAAATGCTATCAATGGGTGCGATGCTACAGAGAAATCTATCATG 312  
 Db 14379 GTTCTGTTGCTTAAATGCTATCAACAGCCGCGATGCTAAAGAGGGATCTTTGTG 14320  
 QY 313 AAAGGAAGAGTCAATTTGATGTGGCAACTTCATTTGTCGTATTTTAAAGAAATTTGTCAG 372  
 Db 14319 AAAGGAAGAGTCAATCTGTGGGCAACTTCATTTGTCGTGTC-----TTAAGAAATTTGCTGT 14265  
 QY 373 GACCAACCCCACTTCAACACCATGACCTGATCATGAGGAGCCATCCACATTTGAGGC 432  
 Db 14264 AGTCACCCCACTTCAAGCAACCATGACCTGATGATGAGGAGCCATCCACATTTGAGGC 14205  
 QY 433 GAGAACCTCCACAG-TAAAGAGATTATGATCTCTAAAGGATCAGATGAACATTAGCAT 491  
 Db 14204 AAGACTATCCACGCAAAAGATTATGATTTCACTGAGGCTCAGATGC---TTGTAT 14149  
 QY 492 TTTTAAAGCAATAAGTATTTTACGTAAGATATGATGTTATTTTATGAGCATATGC 551  
 Db 14148 TTTTAAAGCAATAAGTATTTTATTTTAAAGTATGTACACTGTTTATGATAGAACT 14089  
 QY 552 TATTATGCATTTAATAGACTCCAGTATATTGTAACATAAC-TTTAAATGCACGTGGAGA 610

Search completed: April 12, 2003, 04:33:45  
 Job time : 2685.05 secs

Db 14088 TATTGTACACATGATAGACTACAGTATTTGTGTAATAAACTTTTATATGCACTGGAAA 14029  
 QY 611 TAAAA-----GTATTTGCTCTTTATGATATTTGCTTTTATGCACTAGTCTGTAAT 661  
 Db 14028 CCAAAAACTCATGAGACTAATTTTATGTGATAGTCTTTTATGTCAGCAGTTTGAAC 13969  
 QY 662 GGAACACTACATTTCTCTTTGGGTACACCTGTATACAGAAAGAAA 705  
 Db 13968 TGAACCTGCAATATCTCGTGGTATGCTGTACCCCTGAAAAA 13925

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 14:22:18 ; Search time 1366.46 Seconds  
(without alignments)  
15994.713 Million cell updates/sec

Title: US-09-914-152-3\_COPY\_4000\_4750  
Perfect score: 751  
Sequence: 1 aaagcagactctagcactca.....ggggaggtctacacgttctgc 751

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_in:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vi:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_inv:\*
- 32: em\_htg\_other:\*
- 33: em\_htg\_mus:\*
- 34: em\_htg\_pln:\*
- 35: em\_htg\_rod:\*
- 36: em\_htg\_mam:\*
- 37: em\_htg\_vrt:\*
- 38: em\_sy:\*
- 39: em\_htgo\_hum:\*
- 40: em\_htgo\_mus:\*
- 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	751	100.0	10562	6	E38420 Novel polyp
2	751	100.0	170121	9	AF064860 Homo sapi
3	751	100.0	340000	9	HS21C080 Homo sapi
4	734.2	97.8	185982	2	AC073231 Homo sapi
5	62	8.3	7218	6	I66494 Sequence 14
6	46.4	6.2	196514	2	AC125555 Rattus no
7	43.6	5.8	1161	10	AB012210 Rattus no
8	43.6	5.8	1161	10	AB029611 Rattus no
9	43.6	5.8	1539	6	AR116596 Sequence
10	43.6	5.8	1539	6	AR121155 Sequence
11	43.6	5.8	1539	6	AX454137 Sequence
12	43.6	5.8	1539	10	RNU32673 Sequence
13	43.6	5.8	3122	6	AX087897 Rattus norv
14	43.6	5.8	3122	10	RNU23483 Rattus norv
15	42.2	5.6	212337	2	AC109262 Mus muscu
16	41.2	5.5	166358	2	AC123360 Rattus no
17	40.8	5.4	125020	9	AF429315 Homo sapi
18	40.8	5.4	215126	2	AC083916 Mus muscu
19	40.8	5.4	295352	2	AC112276 Rattus no
20	40.6	5.4	125020	9	AF429315 Homo sapi
21	40	5.3	115666	2	AC105744 Oryza sat
22	40	5.3	191883	2	AC126956 Rattus no
23	39.6	5.3	203598	2	AC090250 Homo sapi
24	39.4	5.2	1703	10	AF441863 Mus muscu
25	39.4	5.2	149296	2	AC128909 Rattus no
26	39.4	5.2	198011	10	AL662887 Mouse DNA
27	39.4	5.2	236659	2	AL808021 Mus muscu
28	39.4	5.2	247874	2	AC103340 Rattus no
29	38.8	5.2	118641	2	AC103152 Rattus no
30	38.8	5.2	158992	2	AC110657 Rattus no
31	38.6	5.1	35050	2	AC006099 Homo sapi
32	38.6	5.1	82500	9	AF352033 Homo sapi
33	38.6	5.1	129130	2	AC013252 Homo sapi
34	38.6	5.1	168239	9	AL391421 Human DNA
35	38.6	5.1	186769	2	AC020617 Mus muscu
36	38.6	5.1	198864	2	AC102044 Mus muscu
37	38.6	5.1	202968	2	AC115752 Mus muscu
38	38.6	5.1	236700	2	AC122397 Mus muscu
39	38.4	5.1	148471	9	AC009332 Homo sapi
40	38.4	5.1	164824	9	AC009264 Homo sapi
41	38.4	5.1	187196	10	AF520420 Mus muscu
42	38.4	5.1	187929	9	AC108510 Homo sapi
43	38.4	5.1	189276	2	AC118663 Homo sapi
44	38.2	5.1	154671	2	AC106318 Rattus no
45	38.2	5.1	159707	9	AC013403 Homo sapi

ALIGNMENTS

RESULT 1  
E38420  
LOCUS E38420  
DEFINITION Novel polypeptide.  
ACCESSION E38420  
VERSION E38420.1 GI:18626994  
KEYWORDS JP 2000245464-A/2.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 10562)  
AUTHORS Narimatsu, H., Isshiki, S., Togayauchi, A. and Sasaki, K.  
TITLE Novel polypeptide  
JOURNAL Patent: JP 2000245464-A 2 12-SEP-2000;

10562 bp DNA linear PAT 31-JAN-2002



Query Match 100.0%; Score 751; DB 9; Length 170121; Best Local Similarity 100.0%; Pred. No. 1.5e-203; Matches 751; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	AAAGCAGACTCTAGCAGCTGCGCAACACCTCCCGGGTTTCCTTGGTCTGGAAT	60						
Db	85914	AAAGCAGACTCTAGCAGCTGCGCAACACCTCCCGGGTTTCCTTGGTCTGGAAT	85973						
QY	61	TCCATCCCTGGTTCCACCTGTTACATCACACCTCCCTCAAGGACCAAGTCGATGC	120						
Db	85974	TCCATCCCTGGTTCCACCTGTTACATCACACCTCCCTCAAGGACCAAGTCGATGC	85933						
QY	121	CAGCTCCTTACAGGGGCTCAGAACTGCTCACCAGCTTCTCTCCACCGAGGGCCACAGCCC	180						
Db	85934	CAGCTCCTTACAGGGGCTCAGAACTGCTCACCAGCTTCTCTCCACCGAGGGCCACAGCCC	85993						
QY	181	CTGAGAGCCCTTGAAGCTGAGTCTTTCCTTGCATACCTCTTCGGGCTCATAGTGG	240						
Db	85994	CTGAGAGCCCTTGAAGCTGAGTCTTTCCTTGCATACCTCTTCGGGCTCATAGTGG	86053						
QY	241	GCTTGGCATTTGCCCTTCACTCCAGATCTCTCCTTTCAGGTCACAGGAGTGCATCTGA	300						
Db	86054	GCTTGGCATTTGCCCTTCACTCCAGATCTCTCCTTTCAGGTCACAGGAGTGCATCTGA	86113						
QY	301	ACTTAACCTTTCCAGACCCCTTCCAGTTTTCAGTCTCTTACAGAGGTGACCTTCTGATT	360						
Db	86114	ACTTAACCTTTCCAGACCCCTTCCAGTTTTCAGTCTCTTACAGAGGTGACCTTCTGATT	86173						
QY	361	CCTTTGTCTGTGCCCTGTAGCCTCAGTCAAGCTTAAAGCAAGTCTCCTCACCTGGC	420						
Db	86174	CCTTTGTCTGTGCCCTGTAGCCTCAGTCAAGCTTAAAGCAAGTCTCCTCACCTGGC	86233						
QY	421	CTGGGGAGAGTCCAGGAGCTGCGAGTGCCTGCGGTAGGATCTGATGCCAGATT	480						
Db	86234	CTGGGGAGAGTCCAGGAGCTGCGAGTGCCTGCGGTAGGATCTGATGCCAGATT	86293						
QY	481	TCCGTTAGAGAGCTTTCCTTATCCTGACGGCTCTAGCTTTGTGTACTTACTTGT	540						
Db	86294	TCCGTTAGAGAGCTTTCCTTATCCTGACGGCTCTAGCTTTGTGTACTTACTTGT	86353						
QY	541	CCACTTTAAATCAAAATATACCCAGCAACCACTGTGCACAGTTCTCTGGGTTTCAAG	600						
Db	86354	CCACTTTAAATCAAAATATACCCAGCAACCACTGTGCACAGTTCTCTGGGTTTCAAG	86413						
QY	601	AGGATGTAGACATACCCCTTGCCTTCAGGCACTATGCGCAGAGGGGGGAGTGACC	660						
Db	86414	AGGATGTAGACATACCCCTTGCCTTCAGGCACTATGCGCAGAGGGGGGAGTGACC	86473						
QY	661	TAGCAGAGGGCGGAGCAGATGGGATACACTCAGGAGGCTGCGAGGCGCAGAG	720						
Db	86474	TAGCAGAGGGCGGAGCAGATGGGATACACTCAGGAGGCTGCGAGGCGCAGAG	86533						
QY	721	GCAGAGGAGAGGGAGGTCTACACGTTCTGC	751						
Db	86534	GCAGAGGAGAGGGAGGTCTACACGTTCTGC	86564						
RESULT 3									
HS21C080 340000 bp DNA linear PRI 24-MAY-2000									
LOCUS									
DEFINITION Homo sapiens chromosome 21 segment HS21C080.									
ACCESSION AL163280 AP001735 BA000005									
VERSION AL163280.2 GI:7717369									
KEYWORDS									
SOURCE human.									
ORGANISM Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
1 (bases 1 to 340000)									
Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,									
Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,									
Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,									
Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,									

TITLE		JOURNAL
Direct Submission		
Submitted (05-MAY-2000) The Chromosome 21 Mapping and Sequencing Consortium: * RIKEN Genomic Sciences Center, Human Genome Research Group * Institute of Molecular Biotechnology, Genome Analysis * Keio University School of Medicine, Dept. of Molecular Biology * GBF, Dept. of Genome Analysis * Max-Planck Institute for Molecular Genetics (addresses see below)		
The Chromosome 21 Mapping and Sequencing Consortium consists of		
* RIKEN Genomic Sciences Center, Human Genome Research Group, * Sagamihara 228-8555, Japan,		
* e.mail: sakaki@gs.riken.go.jp		
* URL: http://hgp.gsc.riken.go.jp/		
and		
* Institute of Molecular Biotechnology, Genome Analysis, * Beutenbergstrasse 11, D-07745 Jena, Germany,		
* e.mail: gscj-submit@genome.imb-jena.de		
* URL: http://genome.imb-jena.de/		
and		
* Keio University School of Medicine, Dept. of Molecular Biology, * Tokyo 160-8582, Japan,		
* e.mail: shimizu@mb-med.keio.ac.jp		
* URL: http://adenine.dmb.med.keio.ac.jp/		
and		
* GBF, Dept. of Genome Analysis,		
* Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e.mail: info.genome@gbf.de		
* URL: http://genome.gbf.de/		
and		
* Max-Planck Institute for Molecular Genetics,		
* Ihnestrasse 73, D-14195 Berlin, Germany,		
* e.mail: info-chr21@molgen.mpg.de		
* URL: http://chr21.rz-berlin.mpg.de/.		
Location/Qualifiers		
1..340000		
/organism="Homo sapiens"		
/db_xref="taxon:9606"		
/chromosome="21"		
/map="21q22.3"		
<1..125946		
/organism="Homo sapiens"		
/db_xref="taxon:9606"		
/chromosome="21"		
/map="21q22.3"		
/clone="P16019, 5' partial"		
/clone_lib="RPC11,3-5 PAC library"		
/note="Accession No. AF121897"		
73410..243533		
/organism="Homo sapiens"		
/db_xref="taxon:9606"		
/chromosome="21"		
/map="21q22.3"		
/clone="P70124"		
/clone_lib="RPC11,3-5 PAC library"		
/note="Accession No. AF064860"		
224137..>340000		
/organism="Homo sapiens"		
/db_xref="taxon:9606"		
/chromosome="21"		
/map="21q22.3"		
/clone="P206A10, 3' partial"		
/clone_lib="RPC11,3-5 PAC library"		
/note="Accession No. AF121782"		
286628..>340000		
/organism="Homo sapiens"		

```
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="BAC-291B3, 3' partial"
/clone_lib="BAC library"
/note="Accession No. AF064857"
1088..1366
/note="LIM4"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
2398..2495
/note="(TCTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
complement(2496..2878)
/note="THELC"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
2879..2931
/note="(TCTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
2932..3004
/note="(TCCA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
3005..3074
/note="(TCCA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
3473..19157
/gene="SH3BGR"
join(<3473..3565,12649..12678,15331..15450,18739..19157)
/gene="SH3BGR"
join(3473..3565,12649..12678,15331..15426)
/partial
/gene="SH3BGR"
/note="Accession No. X93498"
/codon_start=1
/product="21-Glutamic Acid Rich protein 21-GARP"
/protein_id="CA990445.1"
/db_xref="GI:7717370"
/translation="GSEKAEGETEAQKESGSDVGNLPEAQKNEEGEATETEE
TAMEGAEQEAEEETAEQEEPEDEDS"
3473..3565
/gene="SH3BGR"
/number=4
5542..5829
/note="AluX"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(6887..7067)
/note="LIMD3"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
complement(8482..8762)
/note="AluJo"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(8865..8964)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
9131..9424
/note="AluSq"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(9723..9882)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
complement(9989..10077)
/note="L2"
```

```
repeat_region
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
complement(10078..10429)
/note="THEIB"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
complement(10430..10651)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
10980..11000
/note="(TTTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
complement(11002..11282)
/note="AluSp"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(11315..11556)
/note="AluJo"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
11702..11753
/note="MIR"
/rpt_family="SINE/MIR"
/rpt_type=DISPERSED
12649..12678
/gene="SH3BGR"
/number=5
complement(13115..13425)
/note="AluX"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(13598..13892)
/note="AluY"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
14013..14131
/note="(CA)n"
/rpt_family="Simple_repeat"
```

```
Query Match 100.0%; Score 751; DB 9; Length 340000;
Best Local Similarity 100.0%; Pred. No. 1.5e-203;
Matches 751; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAGCAGACTCTAGCAGCTCAGGACTGCCAACACCTCCCGGGTTTCCTTGGTCTGGAAT 60
|||||
Db 159222 AAAGCAGACTCTAGCAGCTCAGGACTGCCAACACCTCCCGGGTTTCCTTGGTCTGGAAT 159281

QY 61 TCCATCCCTGTTCCACCTGTTACATCACACCTCCCTTCAAGGACCAGTGCAGATGC 120
|||||
Db 159282 TCCATCCCTGTTCCACCTGTTACATCACACCTCCCTTCAAGGACCAGTGCAGATGC 159341

QY 121 CAGCTCCTTCAAGGGGCTCAGAAATGCTCACCAGCTTCCTCTCCACGAGGGCCACACCCC 180
|||||
Db 159342 CAGCTCCTTCAAGGGGCTCAGAAATGCTCACCAGCTTCCTCTCCACGAGGGCCACACCCC 159401

QY 181 CTGGAGACCCCTTGAAGTGAAGTCTTGTCTTGTGATCTCTTCTGGCCCTCATAGTGG 240
|||||
Db 159402 CTGGAGACCCCTTGAAGTGAAGTCTTGTCTTGTGATCTCTTCTGGCCCTCATAGTGG 159461

QY 241 GCTTGGCCATTGCTCCCTCACTCCAGATCTCTCTTTCAGGTCACGAGATGATCTTGA 300
|||||
Db 159462 GCTTGGCCATTGCTCCCTCACTCCAGATCTCTCTTTCAGGTCACGAGATGATCTTGA 159521

QY 301 ACTTAAGTCTTCCAGACCCCTTTCAGTCTTCCAGTCTTTCAGTCTTTCAGAGAGTGGACTTCTGATT 360
|||||
Db 159522 ACTTAAGTCTTCCAGACCCCTTTCAGTCTTCCAGTCTTTCAGTCTTTCAGAGAGTGGACTTCTGATT 159581

QY 361 CCTTTGTCTCTGTGCCCTGTAGCCTCAGGTCAAGGCTTAAGGCAAGGTCTCTCCTCACCTGGC 420
|||||
Db 159582 CCTTTGTCTCTGTGCCCTGTAGCCTCAGGTCAAGGCTTAAGGCAAGGTCTCTCCTCACCTGGC 159641
```



[illegible][illegible]